PRICE THEORY

PRICE THEORY

BY

W. J. L. RYAN

FELLOW OF TRINITY COLLEGE, DUBLIN

LONDON
MACMILLAN & CO LTD
NEW YORK · ST MARTIN'S PRESS

MACMILLAN AND COMPANY LIMITED
London Bombay Calcutta Madras Melbourne

THE MACMILLAN COMPANY OF CANADA LIMITED
Toronto

ST MARTIN'S PRESS INC
New York

PRINTED IN GREAT BRITAIN

PREFACE

It is tempting to begin by defining the scope of economics and describing the methods by which economic truths are customarily pursued in academic circles. The temptation is acute for an economist, for the fascination of economics with its own scope and method verges on neurosis. It is with reluctance, therefore, that we do not deal with these topics explicitly. We shall not prejudice their importance, however, if we define economics as the kinds of thing that economists habitually talk about, and its methodology as the way in which they customarily do so.

Economists generally describe certain decisions that are taken by individuals who are acting on their own behalf, or as agents, in a free society, and attempt to explore some of their effects. The kinds of decision that interest economists are those which lead to a purchase or to a sale. In the Western world, those who decide to buy and sell may be classified roughly into households, firms and the various agencies of government. Each household decides what commodities and services to buy and when, where and in what quantities to buy them. These decisions make up the *purchase plan* of the household. Each household will also have a *sales plan* setting out the things that its members have decided to sell and the quantities, prices and places at which they will be sold. The sales and purchase plans of the household will be related to one another, for the sums of money that the members of the household get from selling their labour or lending their savings or renting their land generally constitute the fund out of which they buy the goods and services of everyday consumption.

Similarly, each firm in the economy must decide what goods to produce and sell and when, where and the quantities in which to sell them. All these decisions make up the *sales plan* of the firm. In addition, each firm must decide what things to use in making its products, and when, where, how, and in what quantities to use them. All decisions of this kind are summarised in its *purchase plan.*

The purchase and sales plans of the firms are not independent of one another, for firms buy in order to sell. The sums of money that they earn by selling the goods they produce are used directly or indirectly to pay for the things they require to assist in their production and sale. We would expect, too, some relation between the plans of households and those of firms. The things that firms plan to sell must be similar to those which households plan to buy, and the things that firms plan to buy must be more or less the same as the things that households or other firms are planning to sell.

In a free world the implementation and revision of these plans affect almost all facets of human life and endeavour. As economists, however, we are primarily interested in how these plans determine both relative prices and price levels. As firms and households act on the plans they have made, the relationship between prices may alter: butter may become more expensive than nails or bread less dear as compared with jam. And almost all prices might rise as they have done since 1939, or fall as they did in the early 1930's. These twin effects are inextricably and indistinguishably linked together, but if we are to grasp their nature we must examine each in isolation. In this book we are primarily concerned with the determination of the relationship between the prices of the things that are bought and sold.

This book is intended as a text-book for students who are planning to specialise in economics. I have tried to state all the assumptions explicitly and to keep the analysis rigorous. The analysis may occasionally seem to be a trifle self-conscious, for I believe that it is important for students to learn not only what economists do but why and how they do it. There are frequent summaries of the analyses, and I hope that these will be more helpful than they are tedious. I do not think that there is anything that is original in the contents of this book, but there may be some originality in the form in which they are presented.

In elaborating the theory of relative prices, I have used only the traditional tools of analysis. While these tools are suffering a rapid obsolescence, they still do a better job than the prototypes of the tools which may soon supplant them and which are briefly described in the final chapter. It is not improbable, however, that were this book being written five or ten years

later, the emphasis given to the various tools would have to be completely reversed.

I am deeply indebted, either directly or indirectly, to all economists who have written on the theory of price. If I make no attempt to acknowledge my debts in detail, it is because they are too numerous and because I have forgotten the transactions in which many of them originated. I wish to express my gratitude to Professor G. A. Duncan, Professor A. T. Peacock, Professor G. L. S. Shackle, Dr. A. W. H. Phillips, Mr. Jack Wiseman and Mr. F. P. R. Brechling who read the manuscript and made many valuable suggestions and criticisms, and to the students in the London School of Economics and Political Science and in the University of Dublin who forced me to strive after clarity both in thought and expression.

W. J. L. RYAN

TRINITY COLLEGE
DUBLIN

CONTENTS

CHAPTER 1

The Purchase Plan of the Household

The economic unit for the purchase of the goods and services of everyday consumption is the household. It may contain only a bachelor buying for his own immediate needs, or a family with the wife or husband buying on behalf of all. It may be an orphanage or a boarding school with some official devoting his whole time to acquiring goods for the sustenance of the inmates. Irrespective of the size of the household, some individual must decide what goods to buy and how much to buy of them.

Let us illustrate the content of the household's purchase plan by taking a typical housewife, Mrs. Smith. She plans to make the following purchases during the week that lies ahead of her: 2 pounds of butter at 4s. per pound; 10 loaves of bread at 6d. per loaf; 2 pounds of apples at 1s. 6d. per pound, and a whole host of other items, the total cost of all of them adding up, say, to £7 10s. 5d. Mrs. Smith, of course, need not write out her purchase plan in full. She may not even think it out in full: rather, her plans may only appear in the act of purchase. If lack of time or the usual housewife's worries force her into the latter categories, then we can never know her plans directly. We can only construct the plan from her observed purchases by assuming that she always buys what she wants. The implications of this *pis aller* for the science of economics will be explored later in this volume.

We can gain anonymity for Mrs. Smith (or at least make her more anonymous) and gain generality for our treatment, if we state the purchase plan of the typical household algebraically. We shall use 'p' to represent price, and 'q' to represent quantity, and the subscripts 1, 2, 3, ... n, to designate particular commodities. Thus, $p_x \cdot q_x$ would mean the quantity of good X bought multiplied by the price per unit of good X — in other words, the sum of money that the household plans to spend on buying good

X during the week that lies ahead. The complete purchase plan of the household would then be as follows:

$$p_1 \cdot q_1 + p_2 \cdot q_2 + p_3 \cdot q_3 + \dots p_n \cdot q_n = c$$

where '*c*' is the total sum of money that the household plans to spend on the purchase of all goods and services during the week. By expressing the household's purchase plan in this way, we are not introducing mathematics; rather, we are merely translating a statement in the English language into symbols, in the same way as we might translate it into French or Gaelic. It is convenient for us to do so; it is also fashionable.

Of all the elements of this purchase plan, the prices are generally fixed by the market or by the seller: they are usually beyond the control of any housewife. If Mrs. Smith wants the good, she must pay the price that is asked; if she cannot pay the price she cannot have the good. This has not always been so, nor is it yet the rule in all places. In the early nineteenth century in Western Europe prices were a matter for bargaining between the shopkeeper and the individual customer. In middle eastern bazaars, prices are still a matter for haggling. Even in modern industrial economies, some prices are still arrived at in that kind of way. In this and in the next chapter, despite these exceptions, we shall assume that each household faces given prices for the things that it wants to buy, and that these prices cannot be altered by its own efforts.

The other elements in the purchase plan — '*c*' and the quantities purchased — are within the control of the household. The planned consumption expenditure — '*c*' — for the week ahead will depend, at least, on the following things. First, the income of the household: the more the members of the household earn, the more can the household afford to spend, and vice versa. Whether or not the planned consumption expenditure is related to the income actually earned during the previous week, or to the income that the members of the household expect to earn during the week that lies ahead, is a question of fact. If income is stable from week to week it does not matter which of these we assume, for they will both give the same results; when income is changing from one week to another, we shall assume that the former relationship obtains. We shall view the members of the household

as working from Monday to Friday, and then receiving payment for their work. The housewife's planned purchases for the following week then depend on the sums of money that she receives on Friday. Secondly, each week's income may be supplemented by money that has been left over from the incomes of previous weeks. These bits of income that have not been spent, we shall call *savings*. Or this money may have been bequeathed to the household by late relatives, who accumulated it by refraining from spending all their incomes. No matter where they come from, and no matter how they are held — whether in stockings, or in the Post Office Savings Bank, or in a current account at an ordinary commercial bank — such sums can clearly be used to finance current consumption expenditure in excess of income. Thirdly, the household could get money to spend on the goods and services of everyday consumption by selling assets owned by its members, or by borrowing money on their security. For most households the first of these — wages or salaries or other income — is the most important determinant of planned consumption expenditure. It is generally only in exceptional circumstances that recourse is had to savings, or to the sale or mortgage of houses or furniture or other assets owned by the household. In this chapter we shall assume that the household has already decided how much to spend on buying goods for everyday consumption during the week that lies ahead:* that is, we shall suppose that 'c' is given.

Having decided on the amount that it will spend, and given the prices of the things that it wants to buy, the household must decide what quantity of each good to buy. With any given sum of money and at any given set of prices, a household could buy an infinite number of different combinations of goods and services. For example, if the household planned to buy only two goods, namely, bread and apples, priced at 6d. per loaf and 1s. per pound respectively, and if it planned to spend £1 on buying these, it could buy 40 loaves, or 39 loaves and ½ pound of apples, or 20 loaves and 10 pounds of apples, and so on virtually *ad infinitum*. The only condition to which each combination of bread and apples must conform is that its total cost must add up to £1, the household's planned consumption expenditure. Of all these possible combinations of bread and apples the household

* This decision is described later, in Chapter 6.

must choose one. The combination which the household chooses is its purchase plan for the period in question.

Every household makes this choice. As economists, we cannot explain why a household prefers the chosen combination to all others. We merely accept that in deciding on its purchase plan the household, in the light of the desires of its members, chooses those of the goods that it can buy that will satisfy these wants most fully. We may speak of the chosen combination as being that which maximises the household's satisfaction or welfare or ophelimity. By doing so, however, we are not providing a reason for the household's actions; rather, we are merely saying in a different way that the household chooses that combination of goods and services which it prefers to all others. Economists leave the quest for a reason for a household's choices to other disciplines. They leave also to those whose peculiar concern it is the task of passing judgement on the desirability of the household's choices from the moral, aesthetic, ethical, racial, or any other point of view.

We have now described the purchase plan of a household. It will have the following general form:

$$p_1 \cdot q_1 + p_2 \cdot q_2 + p_3 \cdot q_3 + \dots p_n \cdot q_n = c.$$

On the basis of our general observations of economic life, we have assumed that the prices of all the things which it buys are given to the household. From our knowledge of the economic behaviour of the household, we have assumed that the decision as to the size of '*c*', the planned consumption expenditure, precedes the decision as to the quantities that will be purchased. With the given sum of money '*c*' and at the given prices, the quantities of the different goods that are finally chosen are those that the household prefers to all others. That is all that we as economists can say about the nature of the purchase plan of the typical household *per se*.

In the ordinary course of economic activity, however, the purchase plans of the individual households change from time to time. First, the plans may be revised because the household's income changes and therefore makes possible a change in the planned consumption expenditure. Second, the plans may be revised because the prices of some or all of the things that the members of the household would like to buy change. Third, the plans may be altered because the household changes its mind

about the combination of goods that it prefers even though their prices and its income have remained the same. If we are to explain the course of economic activity and especially why prices change from time to time, we must discover how the purchase plans are altered when such changes take place. Further, a knowledge of why and in what way purchase plans are revised is indispensable if our analysis is to assist administrators in formulating economic and social policies. If social opinion holds that the effects of alcohol on the individual organism and on society are undesirable, the government may feel impelled to induce households to modify their purchase plans so that less alcohol is consumed. It may, of course, seek to do this by prohibiting the manufacture and sale of alcohol; or it may pursue the same aim by methods, such as an excise tax, which place less restriction on individual choice. The latter kind of policy has been adopted in the United Kingdom and to a lesser extent in the Irish Republic. Clearly, these British and Irish policies presuppose some knowledge of how consumers of alcohol will react to an increase in its price. For all these reasons, then, we must try to find out how a household's purchase plan will be revised when any of its elements changes.

To help us in our quest, it is convenient to begin by illustrating graphically the purchase or consumption plan of the household. To keep our graphs within the two dimensions of these pages, we shall suppose that the household lives in a world of only two goods, X and Y, whose prices are p_x and p_y respectively, and that the household plans to spend some sum of money 'c' on the purchase of these goods during the week that lies ahead. In Diagram 1 we draw two straight lines or axes at right angles to one another. On the horizontal axis, we measure quantities of good X and on the vertical axis quantities of good Y. Any point which lies between these lines will represent a combination of a particular quantity of X and of a particular quantity of Y. Thus, the point L represents a bundle made up of a quantity OA of X and OB of Y. Now, each such combination of the goods X and Y possesses a certain degree of attractiveness to the members of the household,* for they would enjoy a certain degree of satis-

* We are implicitly assuming that each combination of X and Y will always be used by the household in such a way as to give the greatest satisfaction to its members.

DIAGRAM I

faction or well-offness by having it and consuming it. Our immediate task is to arrange the infinite number of combinations of *X* and *Y* that can be plotted between these axes according to the degree to which they satisfy the desires of the household. This would be an easy task if there existed a meter for measuring satisfactions — a device that could be clamped to the members of the household to measure precisely the extent of the satisfactions they enjoyed as they contemplated eating or consuming different quantities of *X* and *Y*. But no such meter exists; in its absence we must content ourselves with some cruder method of comparing the attractiveness of different quantities of the two goods to the household.

Even without this meter, however, all comparisons are not impossible. We can see immediately that the household will prefer *L* to *M*, for with the combination represented by *L* the household will have a greater quantity of both goods than with the combination denoted by *M*. We can generalise this conclusion by saying that the household will prefer to the combination *M* any bundle of *X* and *Y* represented by a point lying between due north and due east of *M*. The only exception to this rule would be if one of the goods (or both of them) was a 'dis-good' — that is, something like garbage or floodwater, which the household would prefer to be without. We could, of course, extend our analysis to take account of these dis-goods by supposing that they had negative prices, but there is no incentive to do so for house-

holds do not indulge in economic activity in order to acquire them.

Our ability to compare bundles such as M with bundles such as L does not, however, fully solve our problem. It enables us to compare M with all points to the north or to the east of it, but not with a point such as R. With the combination denoted by R the household has more of X but less of Y than with M. No-one but the household can say whether the increased quantity CE of X compensates for the loss of DF of Y. And with that we reach an impasse. Superficially, it might seem that we could elicit this information from the household by direct questioning, but this is an illusion. It would take an interminable time to extract from even a single household the degree of attractiveness of all the infinite number of possible combinations of goods X and Y. The task would be impossible if we ceased to confine ourselves to only two goods and to a single household. Even if all this were possible, our results would not be conclusive, for some households might have changed their minds about what combinations they prefer while the inquiry was proceeding: at one time they might prefer M to R, at another they might prefer R to M. This solution to our difficulties must therefore be ruled out.

Now the relative attractiveness of different combinations of X and Y could be seen immediately on the diagram if we drew a number of lines each passing through points possessing the same degree of attractiveness to the household. Thus, if combinations M and R are equally attractive to the household, we could join M and R and continue this line in both directions to pass through all other combinations of X and Y that the household assesses equally with M and R. This line is called an *indifference curve*, because it joins all points between which the household would be indifferent: if the members of the household believe that they would feel as well off with the combination M as with the combination R, then it would not matter to them which they had. A line similar to that which joins M and R can be drawn through each other point on Diagram 2. The line passing through L, for example, would pass through all combinations which the household values equally with L. And similarly with each other point. If such lines are drawn, then comparisons of the kind that we found impossible earlier can be easily made. If, for example,

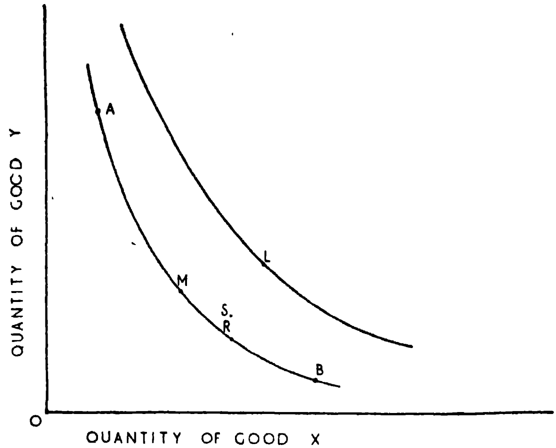


DIAGRAM 2

we wish to compare M with the combination S which represents more of X and less of Y than does M, we can do so by comparing S with R which lies vertically below it on the same indifference curve as M. It is clear that S is preferable to R, for with S the household has as much of X as at R but more of Y. If S is preferable to R and if the household is indifferent as between R and M, then S must be preferable to M.

It should now be clear that the assumption that we are seeking about the order of attractiveness of the bundles of X and Y that are plotted on the diagram must be an assumption about the shape of these indifference curves. Our previous analysis rules out some shapes: an indifference curve cannot rise due northwards or due eastwards or with any degree of north-eastwardness from the origin in the diagram, for such curves would pass through combinations of X and Y between which the household could not possibly be indifferent. It is inconceivable, for example, that M and L should lie on the same indifference curve, for with L the household would have more of both goods than with M. This forces us to conclude that indifference curves must follow some north-westward path across the diagram, such as that followed by the lines I_1, I_2, and I_3 in Diagram 3. We must now choose between I_3 which is a straight line, and I_1 and I_2 which when viewed from the origin are convex and concave respectively.

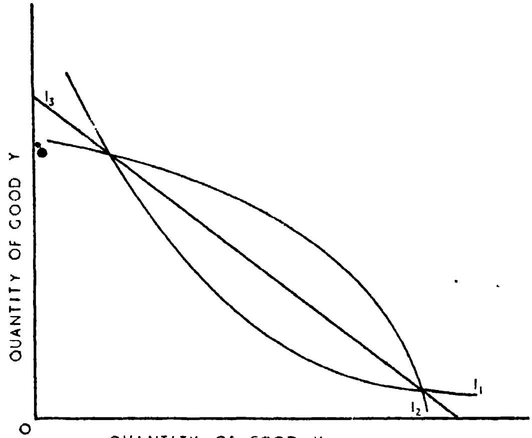


DIAGRAM 3

We choose the shape I_1,* and for two reasons. First, it is in accord with such knowledge of our own preferences as may be obtained by introspection. If I have a small quantity of X and a large quantity of Y — e.g., if I have the combination of X and Y denoted by A in Diagram 2 — I would require a relatively large quantity of Y to compensate me for the loss of one of my units of X; and I would require a greater quantity of Y the smaller the quantity of X which I have. Conversely, if I have the bundle denoted by B in Diagram 2, I would demand more and more X as compensation for the surrender of successive units of Y. In general, the smaller the quantity of X (or Y) and the larger the quantity of Y (or X) which I have, the more highly do I value each unit of X (or Y) and the less is the value I place on each unit of Y (or X).† That is, my indifference curves are convex to the origin. Second, while I cannot prove that the preferences of other households are similar to mine, I suspect that they are, and my suspicions are not inconsistent with other households' behaviour. In the course of economic activity, it is observed that when a good becomes relatively cheaper as compared with other goods more of it is generally bought, and very occasionally less of it may be bought. This behaviour is not inconsistent with

* The implications in a two-good world of the shapes shown by I_2 and I_3 in Diagram 3 are explored in the appendix to this chapter.

† On page 12 *infra*, this notion is stated in terms of the marginal rate of substitution of good X for good Y.

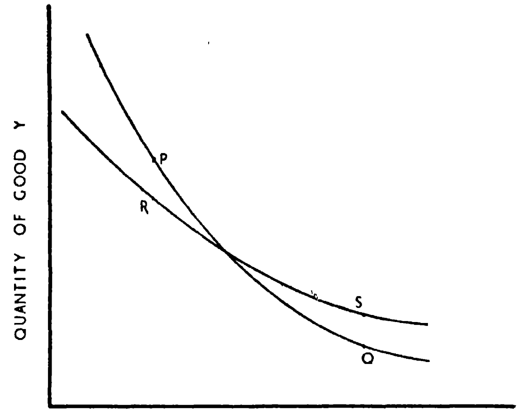


DIAGRAM 4

the hypothesis that the indifference curves of other households are, like mine, convex to the origin.

The assumption that all indifference curves are like I_1 is not, however, sufficiently restrictive, for two indifference curves each of which is convex to the origin might intersect one another. We can show easily that this is impossible. In Diagram 4, the indifference curve that passes through *R* and *S* cuts that which passes through *P* and *Q*. To shorten the proof we take the point *P* to be vertically above *R*, and we let *Q* be vertically below *S*. It can be seen that the household will prefer the bundle of *X* and *Y* represented by *P* to that represented by *R*, for with *P* it has the same quantity of *X* but more of *Y* than with *R*. Since *R* and *S* lie on the same indifference curve the household must be indifferent as between them. The combination of *X* and *Y* represented by *S* must be preferred to that represented by *Q*, for with it the household has more of *Y* than at *Q* but the same quantity of *X*. But if *Q* is less attractive than *S*, and *P* more attractive than *R*, then *P* must be preferred to *Q*, since *R* and *S* are valued equally. But this is impossible since *P* and *Q* lie on the same indifference curve. We assume, then, that all indifference curves are convex to the origin and that they never cut one another.

Each bundle of the goods *X* and *Y* will lie on an indifference curve, and there will be an infinite number of these indifference curves each passing through combinations that are equally

attractive to the household. All these curves make up the household's *indifference map*. The indifference map illustrates the household's tastes or desires for the two goods, and its preferences as between different combinations of them. So long as there is no change in the tastes and preferences of the members of the household the whole indifference map will remain stable. If tastes and preferences change, then the existing indifference map will be replaced by a new one. If, for example, good *Y* is aspirin and good *X* is bread, and if the members of the household develop headaches, then each of the indifference curves will sink towards the horizontal axis, as is shown in Diagram 5: for now that headaches have intensified the desires for aspirin a smaller quantity of aspirin *OF* can be expected to be as attractive to the household as the quantity of bread *OG*. When headaches have been cured, the indifference curves will return to their initial positions.

We must now make clear the implications of our assumption that indifference curves are convex to the origin. This assumption means that as we travel along any indifference curve northwestwards the quantity of *X* that will be given up for each additional unit of *Y* becomes smaller and smaller. In Diagram 6, if the household possesses the bundle of goods *X* and *Y* represented by *G*, by giving up *HG* of *X* for an additional unit *HJ* of *Y*, it would have the bundle *J*, which since it lies on the same indifference curve as *G* is viewed as being equally satisfactory. If the

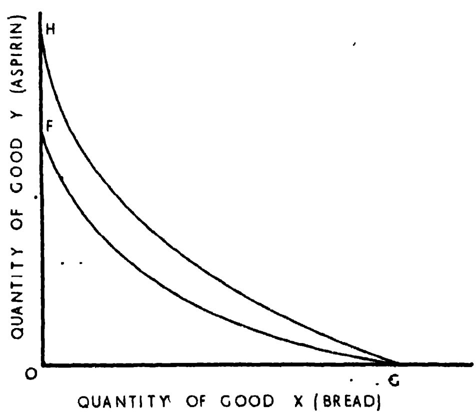


DIAGRAM 5

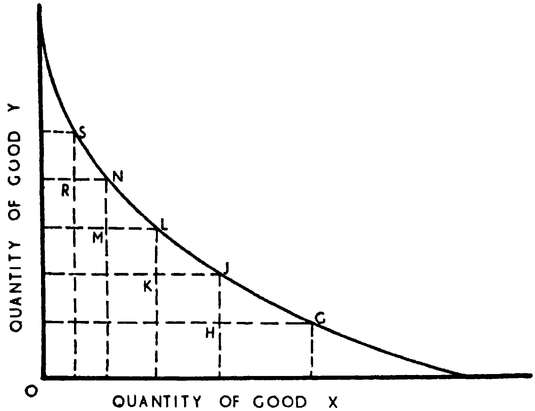


DIAGRAM 6

household possesses *J*, by forgoing *JK* of *X* for the additional unit *KL* of *Y*, it can reach *L* which again is equally attractive with both *G* and *J*. And so on along the whole curve: as the quantity of *Y* is increased, the household would be willing to forgo smaller and smaller quantities of *X* for each additional unit of *Y*.

This property of the typical indifference curve is usually expressed in technical terminology by economists. The quantity of *X* the loss of which in the estimation of the household would just be compensated by an additional unit of *Y* is called the marginal rate of substitution of *X* for *Y*. In Diagram 6 in the range *JG* of the indifference curve, the marginal rate of substitution of *X* for *Y* is equal to GH/HJ. In the range *LN*, the marginal rate of substitution of *X* for *Y* is LM/MN. Our assumption that indifference curves are convex to the origin, when translated into this new terminology, becomes an assumption that in the household's estimation the marginal rate of substitution of good *X* for good *Y* decreases, as more of good *Y* is acquired by surrendering good *X*.

We shall assume — and not unreasonably — that the economic ambition of each household in its role as consumer is to be as well-off as possible — to acquire that bundle of goods and services that promises the fullest satisfactions of its desires. When translated into our jargon, this ambition becomes a striving to reach the highest possible indifference curve. In this pursuit, the household is hampered by the size of the sum of money '*c*' that it plans to spend on buying goods and services and by the

fact that it must pay a price for each unit of each good that it buys. The nature of these limits can be illustrated on the same kind of diagram as that on which we drew the indifference map. On the vertical axis of Diagram 7 we measure quantities of *Y* and on the horizontal axis we measure quantities of *X*. The distance *OM* shows the quantity of *X* that can be bought with the given planned consumption expenditure '*c*' when the price is p_x per unit; *OL* is the quantity of *Y* that the household can buy with the same sum '*c*' when the price p_y has to be paid for each unit of it. The straight line joining the points *L* and *M* will pass through all combinations of goods *X* and *Y* that the household can acquire if it spends all of the sum of money '*c*' on their purchase at the given prices.*

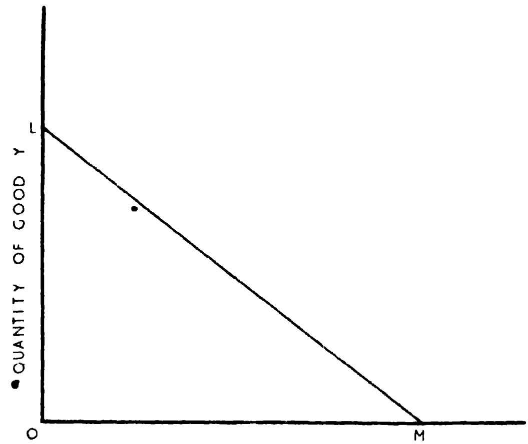


DIAGRAM 7

* This can be proved simply with the help of Diagram 8. Let *T* be any combination of the goods *X* and *Y* lying on the straight line *LM*. In purchasing the bundle of *X* and *Y* that is represented by *T* the household spends $OR \cdot p_y$ plus $OS \cdot p_x$.

Now, $OM = c/p_x$,

and $OL = c/p_y$.

That is: $p_x = c/OM$,

and $p_y = c/OL$.

Substituting these values for p_x and p_y, we get:

Total expenditure $= c\,(OR/OL + OS/OM)$.

But $OR/OL = MT/ML = MS/OM$;

Therefore, total expenditure $= c\,(MS/OM + OS/OM)$

$= c\,(MS + OS)/OM$

$= c.$

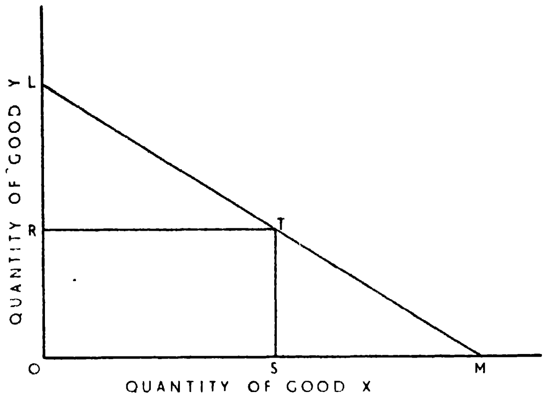


DIAGRAM 8

The making of a purchase plan is merely the choice by the household of *one* of the many purchaseable combinations of the goods *X* and *Y* that lie on the 'budget' line *LM*. By superimposing this budget line on the household's indifference map, we can show which particular combination of *X* and *Y* will be chosen and why.

In Diagram 9, the budget line is drawn between the same axes as the household's indifference map. The household will choose the quantities of *X* and *Y* represented by the point *P*, where the budget line touches an indifference curve. There can

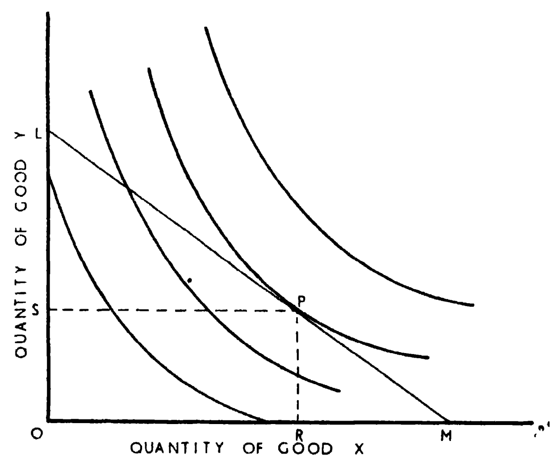


DIAGRAM 9

only be one point such as P since no two indifference curves can cut one another and all are convex to the origin. Any combination other than P can be shown to be less attractive in the estimation of the household: for all other points on the budget line must lie on indifference curves of a lower order than that on which P lies. With its given planned consumption expenditure and at the prices ruling in the market, the household will satisfy its wants most fully by buying the bundle of X and Y represented by the point P. The purchase plan of the household is then:

$$OS \,.\, p_y \text{ plus } OR \,.\, p_x = c.$$

When the household is implementing this plan we say that it is in *equilibrium*, for it is satisfying its desires as fully as the circumstances in which it finds itself permit.

A property of the equilibrium purchase plan of the household is that the marginal rate of substitution of good X for good Y is equal to the ratio of their prices. In Diagram 9, the marginal rate of substitution of Y for X when the household is planning to buy the bundle P is equal to the slope of the indifference curve at P, which is the same as the slope of the budget line LM. The slope of LM is equal to OL/OM. We know, however, that OM equals c/p_x and that OL equals c/p_y — that is, the slope of LM is equal to c/p_y divided by c/p_x, or p_x/p_y. To say that the household chooses that purchase plan which equates the marginal rate of substitution with the ratio of the prices is merely to say in a different way that it chooses that which promises to maximise its satisfactions from consumption. We must not speak of the former statement as if it were the reason for the latter. Nor should we ever speak of the household as trying to equate the marginal rate of substitution with the price-ratio, for from that it is too easy to slide into a behaviouristic interpretation of the indifference curves and budget line — to view the flesh-and-blood household as straddling along an indifference curve and budget line until its feet come together. The household seeks to satisfy its desires as fully as its limited expenditure and prices permit; when it is succeeding in doing this, its purchase plan may be illustrated on the diagram by certain lines. Some of the lines on the diagram bear a certain relationship to one another, which may be expressed geometrically or algebraically. These rela-

tionships are not the reasons for the choice of the purchase plan, nor can they tell us anything about the methods by which the household chooses it. We may illustrate these distinctions by an analogy. A man living in London plans to visit Bristol, which lies 110 miles due west of London. When his journey has been completed, we describe his position as being in Bristol. We add nothing of substance to this by saying that he is now 110 miles due west of London. And if we do say this, we do not imply that he followed a compass course in journeying to Bristol. This point has been laboured because it is one that recurs in all branches of economic analysis. If we neglect it, we risk appearing to re-create the real world in the image of our analysis.

We have now succeeded in illustrating diagrammatically a household's purchase plan — the same plan that we illustrated arithmetically on page 1 and expressed algebraically on page 2. We have added nothing to our knowledge by doing so, but we have presented what knowledge we had in such a way that we can easily deduce from it how the purchase plan will be revised if circumstances should change. It should be clear from Diagram 9 that the purchase plan will be revised if there is any change in one or more of the prices of the things that it buys, or a change in its tastes and preferences, or any alteration in the household's planned consumption expenditure, or any change in any two or in all three of these. We shall examine the nature of these revisions in turn.

First, the reaction of consumption to changes in planned expenditure. Let us take a numerical example. Initially, we shall suppose that the planned expenditure of the household is 50s. and that the prices of *X* and *Y* are 5s. and 10s. respectively. The purchase plan of the household will then be represented by the point *P* in Diagram 10.* If the planned expenditure of the household had been 100s., it would have planned to purchase the bundle of goods *X* and *Y* represented by the point P_1, its tastes and preferences and the prices of *X* and *Y* remaining the same. If its planned expenditure had been 150s., its purchase plan would have been represented by the point P_2; if 200s., by P_3, and so on. There will be a separate purchase plan for each level of planned expenditure.

When the points P_1, P_2, P_3, etc., ar[redacted]gether we get

* For visual reasons, the figures in Diagram 1[redacted]naccurately.

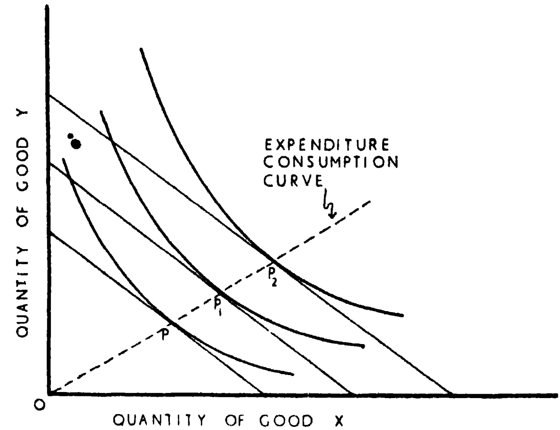


DIAGRAM 10 (*a*)

the *expenditure-consumption curve* of the household. This curve shows us the quantities of the two goods that the household would plan to buy at different levels of expenditure, if its tastes and preferences and the prices of these goods had remained the same. The figures in Diagram 10 show us that our assumption that all indifference curves are convex when viewed from the origin does not uniquely determine the shape of this curve. In Diagram 10(*a*), the expenditure-consumption curve slopes upwards. Our general knowledge of how households react to an increase in their incomes or in their wealth suggests that most expenditure-consumption curves are of this shape, for the increase in expenditure is usually distributed over most of the goods that the household buys. In Diagram 10(*b*), the expenditure-consumption curve begins to move towards the axis on which we measure *Y*, showing that after a certain point, as expenditure rises less of *X* is bought. In Diagram 10(*c*), the curve curls towards the axis on which we measure *X*, showing that as expenditure increases ultimately less of *Y* is bought. Expenditure-consumption curves of these shapes, while rare, are not unknown. Many economists have observed that when for any reason a poor household is enabled to spend more, it may buy less margarine or fewer potatoes, or a smaller number of loaves. It may choose to satisfy its hunger with goods that are more palatable and less monotonous, such as butter, vegetables, fruit and cake. Those goods of which the quantity that the household

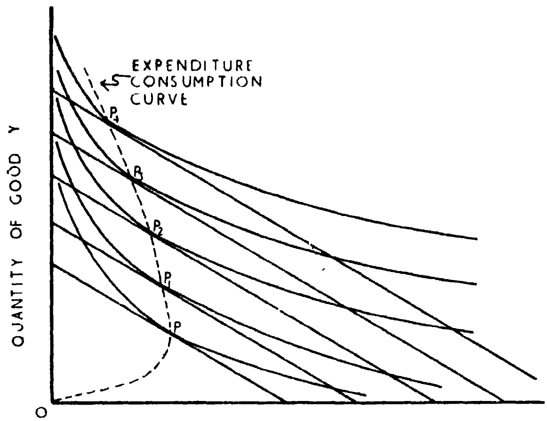


DIAGRAM 10 (*b*)

plans to purchase falls as planned expenditure rises are called *inferior goods*: thus, in Diagram 10(*b*) good *X* and in Diagram 10(*c*) good *Y* are inferior goods.

Diagram 10 shows us that an indifference map consisting of indifference curves each of which is convex to the origin can illustrate both types of reaction to a change in planned consumption expenditure. Indifference analysis can never tell us why a good is inferior. Knowing that it is inferior, however, we can illustrate how less may be bought as expenditure rises by drawing our indifference curves with shapes and in positions which

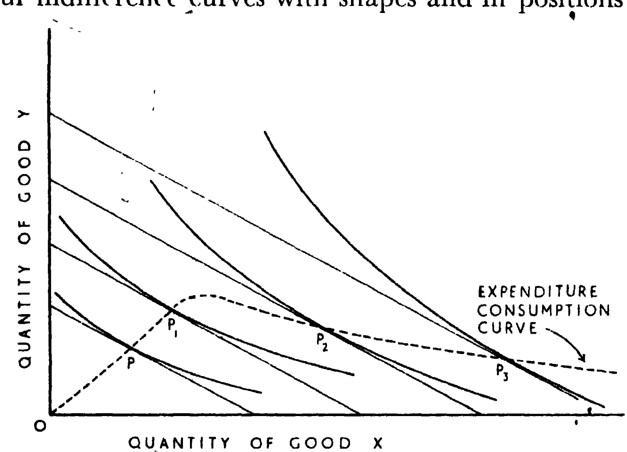


DIAGRAM 10 (*c*)

make this possible, as in 10(*b*) and 10(*c*). Indifference analysis describes; it never explains. And it is our knowledge of the economic behaviour of actual households that tells us what to describe.

Next, we shall describe the revision of a purchase plan when the price of one of the goods that the household plans to buy changes. Initially, suppose that the planned expenditure of the household is 100s. and that the prices of *X* and *Y* are 10s. and 5s. respectively. The purchase plan of the household in these circumstances is then represented by the point *P* in Diagram 11. If the price of *X* had been 7s. 6d. and not 10s. then, all other things remaining the same, $13\frac{1}{3}$ and not 10 units of it could have been bought, and the budget line would have been *LN* and not *LM*, and P_1 and not *P* would have been the purchase plan. If the price of *X* had been 5s., the purchase plan would have been represented by the point P_2. In the same way we can discover what the purchase plan would have been for each other price at which *X* might be sold. When all the points such as *P*, P_1, P_2, etc., are joined together, we have the *price-consumption curve* of the household for good *X*. This shows us how the purchase plan of the household would be revised if the price of *X* changed, the household's tastes and preferences, and its planned consumption expenditure and the price of *Y* remaining the same. Here, again, our assumption that all indifference curves are convex to the origin does not uniquely determine the shape of the price-

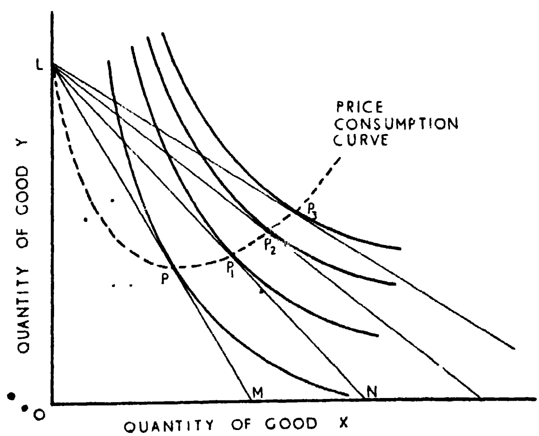


DIAGRAM 11 (*a*)

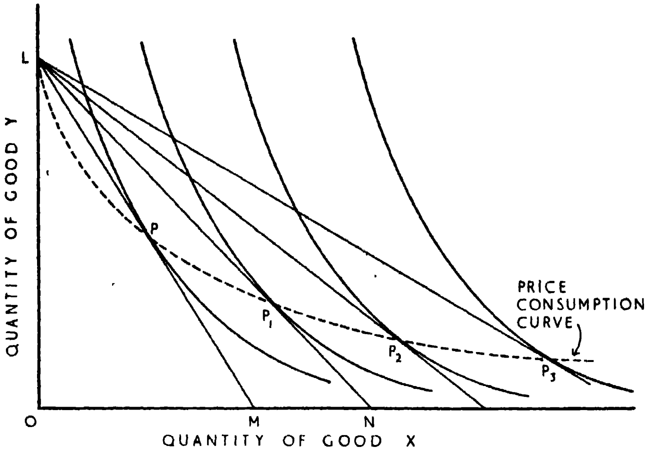


DIAGRAM 11 (*b*)

consumption curve. The different possibilities are classified in Diagram 11. The first two figures — 11(*a*) and 11(*b*) — show the household planning to increase its purchases of *X* when its price falls. This illustrates the most common reaction to a price reduction, for we know that most households buy more of a good when it becomes relatively cheaper. The price-consumption curve shown in Diagram 11(*c*), while not common, is possible. There, when the price of *X* falls below 7s. 6d. per unit, less and less of *X* is bought. When households react in this way —

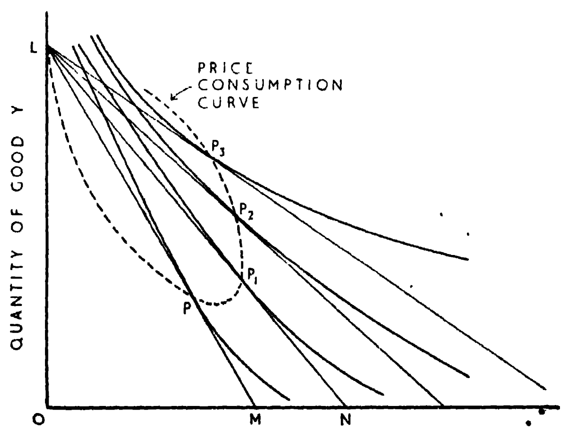


DIAGRAM 11 (*c*)

when they buy less of a good as it becomes relatively cheaper — we refer to that good as a 'Giffen' good.*

This 'inferior' reaction of households to a relative fall in the prices of such goods as margarine and potatoes may be described in a different way. In all the figures that make up Diagram 11, the household is better-off the lower is the price of X, for the plans illustrated by P_1, P_2, P_3, etc., lie on progressively higher indifference curves. P_1 is preferred to P, P_2 to P_1, and so on. We may think of the rejection of plan P and the adoption of plan P_1, as a result of the fall in the price of X from 10s. to 7s. 6d. per unit, as a 'movement' along the price-consumption curve for X, and we shall regard this relative cheapening of X as the 'force' which pushes the household in this direction. This 'force' can be thought of as being the resultant of two other forces. The first of these components is the feeling of betteroffness that a household experiences when the price of even one of the things that it buys falls. With its given planned consumption expenditure, it could now buy the same quantity of each good as it did before the fall in the price of one of them, and have some money left over. It is as if all prices had remained unchanged and the household had been enabled to increase its planned expenditure. Clearly, for any given fall in the price of any good X, the size of this 'gain' will be the greater the larger does X bulk in the household's purchase plan. We can therefore think of this first force as operating along the expenditure-consumption curve. It is called the *income effect*, because the increase in the potential purchasing power of the household's planned expenditure that follows a relative fall in the price of one of the goods that it buys, is as if its income had risen and all prices had remained at the same levels. When its power to acquire goods rises for this reason, the household must decide how this power is to be exercised: it must decide how the increase in its purchasing power is to be spread over the different goods that it buys. This re-allocation of the household's expenditure is the second component of the movement from P to P_1. The household will always re-allocate its expenditure in such a way as to buy relatively more of the good X whose price has fallen, and

* So named after Sir Robert Giffen who is alleged to have observed that when the price of bread rose the poor bought more bread and less meat and less of some other more expensive foodstuffs.

less of good Y. This second force is called the *substitution effect*, and can be viewed as operating along the indifference curve on which P_1 lies.

These notional components of the movement from P to P_1 are shown explicitly in Diagram 12.* LM is the budget line when the household's planned expenditure is 100s. and when the prices of X and Y are 10s. and 5s. respectively. The purchase plan of the household is then represented by P. When the price of X falls to 7s. 6d. per unit, LW becomes the new budget line, and the household rejects P and chooses the purchase plan represented by P_1. RS is a hypothetical budget line showing the different quantities of X and Y that the household could buy if the price of X had remained the same at 10s. per unit and planned expenditure had been increased by an amount just sufficient to leave the household feeling potentially as well-off as it does when expenditure is 100s. and the price of X has fallen to 7s. 6d. per unit. RS touches the higher indifference curve at T; the line passing through P and T is then the expenditure-consumption curve of the household. The line PT illustrates both the magnitude and direction of the income effect. The portion TP_1 of the higher indifference curve illustrates the direction and magnitude of the substitution effect. The substitution effect always means

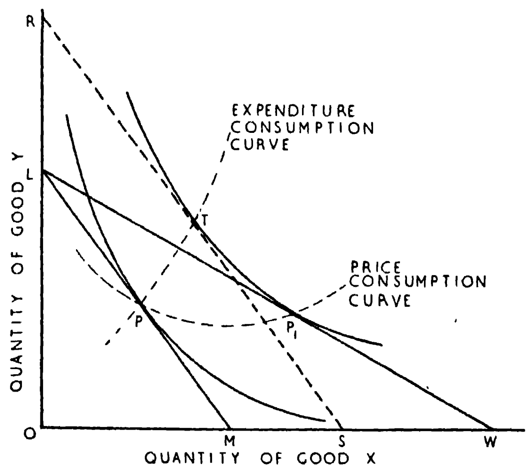


DIAGRAM 12

* For visual reasons, the figure in Diagram 12 is not an accurate portrayal of the following numerical example.

that X, the good that is now relatively cheaper, is substituted for Y, which is now relatively dearer. In other words, the substitution effect on the quantity of X that is bought is always positive. The income effect on X may be either positive or negative, as can be seen from Diagram 10(a) and 10(c) respectively. In Diagram 11(b), the income effect is positive and reinforces the substitution effect in increasing the quantity of X that is bought when there is a relative fall in its price. In Diagram 11(a), the income effect is negative but its influence in tending to reduce the quantity of X that is bought is more than overcome by the positive substitution effect. In Diagram 11(c), the income effect is so negative that it more than offsets the substitution effect, so that less of X is bought in the new purchase plan P_2 than in P_1, the old one. We have now discovered an alternative way of describing a Giffen good: a good is a Giffen good when the negative income effect of a relative fall in its price is stronger than the positive substitution effect.

We must be clear that we have found merely an alternative way in which economists may speak to one another about a Giffen good. In segregating the income and substitution effects, we have not discovered, or said, anything about the mental processes of any flesh-and-blood household. As economists, the only facts we have are that generally households react to a relative fall in the price of a good by buying more of it, but that sometimes they buy less of it. No household has an explicit indifference map, and no household consciously segregates income and substitution effects. Their observed behaviour, however, is *as if* they had the former and experienced the latter. In this respect the theory of consumption in economics is similar to certain branches of physics. Professor O. R. Frisch says in the *Listener* (January 14): 'The theoretical physicist can do amazing things: for instance, from nothing but the number of electrons in an atom of some chemical element he can work out many properties, physical and chemical, of that element; . . . But before he can begin to calculate he must assume that electrons have certain properties; why they have them and why they are all alike, that he cannot explain, as yet.' Similarly, from the indifference maps of households the theoretical economist can work out how they are likely to react to changes in their wealth or income, or in the relative prices of the things they buy. But

before he can begin to do this, he must assume that indifference curves have certain properties; why they have them and why they are all alike, that he cannot explain, as yet.

So far, we have only illustrated the direction in which purchase plans will be revised when there are changes in the things on which they are based. In explaining the behaviour of relative prices in a modern economy, the speed with which plans are revised is as important as the direction in which they are revised. Flesh-and-blood households, and those who buy on their behalf, are frequently the slaves of habit. If a household has for some time been implementing the weekly purchase plan represented by P in Diagram 10, it may be that it will continue for some time to do so when its income or wealth increases. During this period, while revision is pending, the amount of the increase in household income will be added to savings each week. From the purely analytical point of view, we could, of course, think of the maintenance of the plan P as being the result of a change in the household's tastes and preferences, such that the new and higher budget line made possible by the increase in income touched an indifference curve in the new indifference map at P. We could then think of the ultimate rejection of P and the choice of P_1 as being the result of a further change in tastes and preferences. This, however, would be rather a laborious way of illustrating what happens. It is simpler to think of the household as continuing to implement a purchase plan which is not that which maximises its satisfactions, because of laziness, or habit, or the pressure of other decisions. Ultimately, though at the household's pleasure, the revision takes place. All that we can say is that the revision will tend to take place the quicker, the 'poorer' is the household. If a household's income is just sufficient to enable it to satisfy its most basic wants, then any increase in income is likely to be used almost immediately to buy more of the things that the household wants.

These same considerations may explain the speed with which the revisions illustrated in Diagram 11 take place. The household may maintain the plan P when the price of X falls, and temporarily add the unspent balance of its income to savings. As we shall see presently and in later chapters, the effects of households' decisions in determining prices will only occur when the purchase plans are actually revised.

THE DEMAND FOR A COMMODITY

We began this chapter knowing what a purchase plan was, and knowing also, from our general observations of households' behaviour, how these plans might be revised when either the income of the household or the relation between the prices of the things that it bought changed. Thus far, we have been trying to present this knowledge in a way that will help us in our main task, namely, that of isolating the role that these revisions of households' plans play in determining the relations between the prices of the things they buy.

The next step in this direction must be to translate our hypothetical household into a context that reflects more closely that of everyday economic life. We left the household in a hypothetical world where only two goods, *X* and *Y*, were available for its consumption. But flesh-and-blood households live in a world where an almost infinite variety of goods and services is offered for their enjoyment. We could begin our quest for realism by introducing a third good, *Z*, and proceed by the same steps and by asking the same questions as in the previous pages. We could draw three axes, each at right angles to one another, on which we measured quantities of *X*, *Y* and *Z* respectively. In the three dimensional space bounded by these axes, every conceivable combination of *X*, *Y* and *Z* could be represented by a point. The household's preferences for different combinations could be illustrated by indifference surfaces. Since the behaviour that we would be trying to explain is the same as it was when only two goods competed for the household's patronage — namely, that when there is a relative fall in the price of any good, households generally buy more of it and very occasionally less of it — our indifference surfaces must be such as to portray this kind of reaction. That is, each indifference surface must be convex when viewed from the origin. We could show by the same kind of *reductio ad absurdum* that indifference surfaces could never intersect one another. The combinations that the household could purchase with any given planned consumption expenditure at any given set of prices for *X*, *Y* and *Z*, could be illustrated by a budget surface. The content of the household's purchase plan would be represented by a point where the budget surface just touched an indifference surface. If we did this,

the only difference between this chapter and the last one would be that its language would be more cumbersome and its concepts and conclusions less susceptible to diagrammatic illustration. Having described a three-good world, we might introduce a fourth good, and then a fifth, and so on, as we gradually approximated towards reality. Fortunately, there is a shortcut that makes most of this toil and trouble unnecessary. But the snail sees much that the kangaroo misses: by taking this shortcut we shall miss something also, notably, a precise appreciation of the idea of complementarity.*

Let us try to measure all the goods that a household might buy except X — the one in which we are interested — on a vertical axis, and the quantities of X on the horizontal axis. The quantities of X can be measured in the appropriate units: in pounds avoirdupois if it is butter, in pints if it is milk. But in what units shall we measure the assortment of all goods other than X? Pounds weight would not be universally applicable, for domestic service has no weight, and the weight of her who provides it is scarcely of economic significance. The only units to which all can be reduced is value: if all goods are valued in terms of some good, then the total value of any assortment can be got by simply adding together the value of each good expressed in terms of the good that is acting as numeraire. In modern economies, money is the good in which all others are valued, so on our vertical axis we shall measure units of money. That is, we shall measure the power to acquire all goods other than X, rather than attempt to measure different quantities and assortments of these goods directly.† If we do this, we can proceed precisely as we did when dealing with a two-good world and order the household's tastes and preferences for different combinations of X and all other goods — that is, of X and money, the power to acquire all these other goods — by drawing indifference curves. As before, if our deductions from the diagram are

* Compare J. R. Hicks, *Value and Capital*, 2nd edition, Chapter 3. We define complementarity later on page 41.

† Strictly, if we wish to identify all goods other than X with money we must assume that the relative prices of these goods and the relative quantities in which they are bought remain constant. On these assumptions, the amount of money spent on goods other than X would be a measure of the aggregate quantity bought. If relative prices and quantities do not remain constant, there is not only no measure of quantity, but there is no precise meaning that can be attached to the concept of aggregate quantity.

to conform with our observations of households' behaviour, we must draw the indifference curves convex to the origin. If we are given the household's planned consumption expenditure, and the price of X and the prices of each of the other goods that the household might buy, we can draw a budget line, showing all the different combinations of X and of all other goods that the household might buy with the sum of money at its disposal. That combination which the household will choose is represented by the point P where this line touches an indifference curve. In implementing this plan, as we can see from Diagram 13, it will spend LR of its planned expenditure on OS of X, and the remainder OR in buying all other goods and services.

By choosing alternative hypothetical prices for X, and assuming that the planned consumption expenditure, tastes and preferences, and the prices of all other goods, remain the same, we can find by the same method the purchase plan appropriate to each of the prices we have taken for X. When the points representing these alternative plans are joined together, we have the price-consumption curve for X. If X is not a Giffen good, this curve will have the shape it has in either Diagram 11 (*a*) or Diagram 11 (*b*); if X is a Giffen good, the curve will have the shape shown in Diagram 11 (*c*).

The price-consumption curve shows us the purchase plan that the household would choose at each price of X, with its

DIAGRAM 13

given planned consumption expenditure and pattern of tastes and preferences, and at the given prices that are being charged for all goods other than X. If we take out of each of these possible plans that part which describes the household's planned purchases of X, we shall have a relationship between the quantities of X that the household plans to buy and the price of X. This relationship is called the *demand for* X, and it is isolated in columns 8 and 9 of Table 1.

The rows in Table 1 describe the content of the purchase plans that are illustrated in Diagram 13 by the points P, P_1 and P_2 respectively. We use 'p' to represent price, and the suffices 1, 2, 3, ... x, to represent particular goods. The quantity purchased is represented by 'q' in plan P, by 'r' in plan P_1, and by 's' in plan P_2. To keep the table manageable in size, we have included only three of the goods other than X that the household buys. Since we have assumed that there are no changes in the household's tastes and preferences, or in the planned consumption expenditure, or in the price of any good other than X, the different contents of the plans P, P_1 and P_2 must be due solely to the changes in the price of X. The relationship that obtains in these circumstances between the price of X and the planned purchases of X per period, and which we have called the demand for X, is shown in columns 8 and 9. The figures in each row of these columns are graphed in Diagram 14, where we

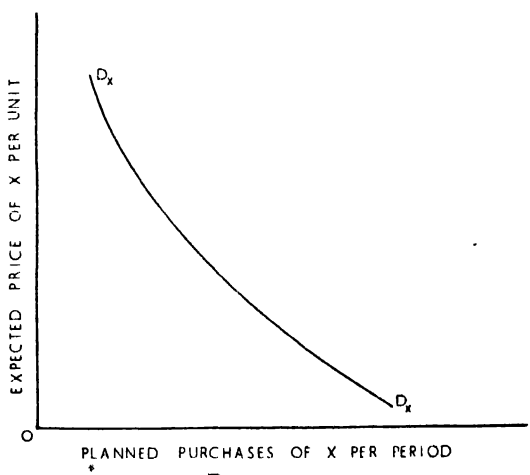


Diagram 14

TABLE I

Alternative purchase plans (1)	*Price of good* 1 (2)	*Planned purchases of good* 1 (3)	*Price of good* 2 (4)	*Planned purchases of good* 2 (5)	*Price of good* 3 (6)	*Planned purchases of good* 3 (7)	*Price of good* X (8)	*Planned purchases of good* X (9)	*Planned consumption expenditure* (10)
Plan P	p_1	q_1	p_2	q_2	p_3	q_3	p_x	q_x	c
Plan P_1	p_1	r_1	p_2	r_2	p_3	r_3	p'_x	r_x	c
Plan P_2	p_1	s_1	p_2	s_2	p_3	s_3	p''_x	s_x	c

measure the price of X on the vertical axis, and the planned purchases of X per period on the horizontal axis. When these points are joined together, we have the *demand curve* for X of the individual household. When X is not a Giffen good, the household's demand curve for it will slope downwards to the right throughout its length. If X is a Giffen good, the demand curve for it will slope downwards to the *left* over the range of prices in which the planned purchases of X rise as its price rises, for we defined a Giffen good as being such that the household's demand for it behaved in that way.

The household whose demand for X is illustrated in Diagram 14 is not the only purchaser of X. We can, however, derive the demand for X of each other household that is a potential purchaser of X in precisely the same way. The *total* or *market* demand for X is obtained by adding together the demands for X of all the households that are planning to buy it. The way in which this summation is effected is illustrated in Diagram 15. Figures A, B and C show the demand curves for X of three separate and independent households. We get the total demand curve for X by adding together the quantities of X that each household plans to buy at each price. Thus at the price p_1, household A plans to buy a_1 of X, B plans to buy b_1, and C plans to buy c_1. The total quantity of X that all households plan to buy at the price p_1 is therefore a_1 plus b_1 plus c_1, and this quantity is plotted against the price p_1 in the figure on the extreme right of the diagram. In the same fashion, we can discover the quantity of X that will be demanded by all households at each other price of X. When all these points are joined we have the total or market demand curve for X.

It is very rarely that a total demand curve will not slope downwards and monotonically to the right. A good may be a

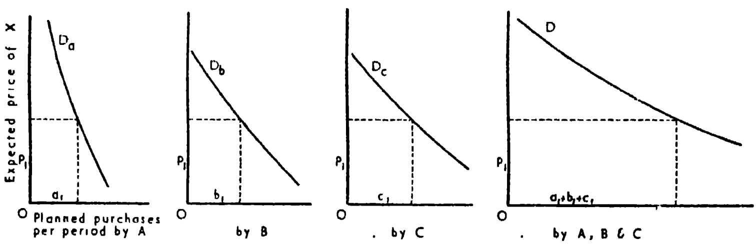


DIAGRAM 15

Giffen good for an individual household, but it is seldom that any given good will be a Giffen good for all households. And even if it is, it is unlikely that it will be so for each household in the same range of prices. In both these cases, there will generally be enough households who increase their planned purchases as the price falls to compensate for those households who buy less because for them the good is a Giffen good in that range of prices.

We have assumed that each household in laying its purchase plan faces given prices for the goods that it may plan to buy. In doing this, we have not been unrealistic, for in any period each household's purchases of any good constitute only a very small proportion of the total quantity of that good that is currently being bought by all households. While each household plans on the assumption that each price is given and beyond its control, the total effect of all households implementing their purchase plans is to assist in the determination of the relations between the prices of the things that they buy. The price-determining role of the purchase plans of households is summarised in the total demand curve for each of the goods that they buy. The analysis of the precise way in which this role is played must be postponed until Chapter 4, for first we must describe the sales plans of firms for each good.

CHANGES IN DEMAND

We began this chapter by describing how the purchase plan of a household depended on its tastes and preferences, on the price at which each good might be bought, and on its planned consumption expenditure. We then described the direction in which the contents of the purchase plan would change if there was some change in planned expenditure, or in the price of one of the goods that the household was planning to buy. Next, we translated the household into a world in which many goods competed for its patronage, and then examined how its purchase plan would alter if there was a change in the price of one of the goods that it bought. From the revised plans we isolated the demand for the good whose price had changed, and we illustrated this operation in Table 1. We have not attempted to explain why the contents of any purchase plan are precisely what they are, and in economics as a positive (as opposed to a

normative) science we never need to do so. We have always been concerned with attempting to explain how the contents of a household's plan will change if there is some alteration in the circumstances in which it is operating. And since our objective in studying economics is to acquire a body of analysis that will enable us to forecast and understand economic events, that is all that we need ever try to do.

We must now examine what will happen to the relationship that we called the demand for X when there is any change in one or other of its determinants — that is, if the household's tastes and preferences, or its planned consumption expenditure, or the price of any other good, should change. In doing so, we shall use Diagram 10, and suppose that on the vertical axis we measure, not the quantity of good Y, but rather money. Given tastes and preferences and the prices of X and of all other goods, the figures in Diagram 10 show how the quantity of X that is purchased changes when there is an increase in planned expenditure. If X is not an inferior good, then more of X will be bought at its given price when the level of expenditure rises; if X is an inferior good, then less of X will be bought at its existing price when expenditure rises, and this we can see from Diagram 10(c). By a diagram similar to Diagram 10, we could measure the effects on the quantity of X purchased of the same increase in planned expenditure at each other price at which X might be bought, on the assumptions that tastes and preferences and the prices of all goods other than X remain the same. We would generally find that the increase in planned expenditure would induce an increase in the quantity of X that the household would plan to purchase at each price. In other words, as a result of the increase in planned expenditure, the whole demand curve for X would move to a new position north-eastwards of its previous position. If there was a reduction in the level of planned expenditure, and if X is not an inferior good, the demand curve for it would move to a new position nearer to the origin. If X were an inferior good, the demand curve would move north-eastwards for a decrease in planned expenditure, and south-westwards for an increase in expenditure.

Next, the effects on the demand for X of a change in the price of one of the other goods that the household is planning to buy. If the price of one of these other goods, Y, falls, then, *ceteris*

paribus, the demand curve for X will move nearer to the origin if Y is a close substitute for X: for in terms of our indifference analysis the positive substitution effect will be most important. If Y is a good that bulks large in the consumer's total spending, then the income effect of the fall in its price may shift the demand curve for X north-eastwards. If X and Y are complements — that is, if they are goods that are customarily used or consumed together, like gin and tonic, or bacon and eggs — then a fall in the price of Y which leads to more of it being bought will cause a rise in the purchases of X, even though its price has remained unaltered. That is, where X and Y are complementary to one another, a fall in the price of Y will cause a rightward shift in the whole demand curve for X. These conclusions are merely stated: the reader may confirm them by the same kinds of argument as were used in the previous paragraph.

Lastly, the effects of a change in the household's tastes and preferences. If the household's preferences for X become stronger relative to its preferences for other goods and services, then, *ceteris paribus*, the demand curve for X will shift north-eastwards, that is, to the right, for households will now plan to buy more of it at each price than before. If there is a relative weakening in the preferences for X, the demand curve for X will sink south-westwards towards the origin — that is, will move to the left.

To sum up: the demand for X is the relationship that obtains between the planned purchases of X and the price of X, when the household's tastes and preferences, its planned consumption expenditure and the price of each other good that it might buy are given. If any one of these determinants of the demand for X changes, the whole relationship that we call the demand for X will change also: that is, the demand curve for X will shift to a new position. If we know whether or not X is an inferior good or whether it is complementary to other goods whose prices may have altered, the direction of the movement in the demand curve for X can generally be predicted. All that we have said about changes in the individual household's demand for X and their causes applies equally to changes in the total or market demand for X.

ELASTICITIES OF DEMAND

Each household must decide, in the light of its tastes and preferences, what quantities of each good to buy at its going price,

with its planned consumption expenditure. These decisions are summarised in the household's purchase plan for the appropriate period of time. So far, we have concentrated on describing how the quantity purchased of each good will vary when the whole purchase plan is revised in response to some change in the pattern of prices or in the level of consumption expenditure. We have divided all commodities and services into two classes: inferior and Giffen goods, and other goods. A good is a Giffen good if more of it is purchased when there is a relative rise in its price, and it is an inferior good if more is purchased when there is a fall in the household's planned expenditure, and vice versa. A good is neither an inferior good nor a Giffen good if more of it is purchased when there is a relative fall in its price or when planned expenditure is increased, and vice versa. The same good will not, of course, always fall into the same class. Margarine, for example, may be neither an inferior nor Giffen good for all households; if it is either for a particular household, it may only be so for some ranges of its price or of the household's planned expenditure.

The goods in each of these classes can be further classified according to the extent to which the quantities of them that are purchased vary as a result of a change in relative prices or in planned expenditure. This responsiveness of the quantity demanded of any good to a change in the data on which the purchase plan is based is called *elasticity*. The responsiveness of the quantity demanded of any good X to a change in its own price, is called the *price elasticity of demand* for X. By the *income elasticity of demand* for any good X we mean the responsiveness of the quantity of it that is demanded to a change in planned consumption expenditure. The responsiveness of the quantity of X that is demanded to a change in the price of some other good Y is called a *cross elasticity of demand* for X. We shall now attempt to measure each of these elasticities in turn.

The way in which the quantity of X that is demanded responds to changes in the price of X is illustrated by the demand curve for it. The strength of the response of the quantity demanded to a relatively small change in the price of X is measured by dividing the proportionate change in the quantity demanded by the small proportionate change in price that causes the change in the quantity demanded. In Diagram 16, when the

households' purchase plans, like bacon and eggs, cheese and biscuits, cigarettes and matches, and gin and lime. If the price of bacon falls, more eggs may be bought to go with the greater quantity of bacon that households are now buying; if gin becomes cheaper, more lime juice may be bought to dilute the larger quantity of gin that the households are now drinking. Goods between which such a relation holds are called complementary goods.

While we shall generally find that goods as between which the cross elasticity of demand is positive are goods which would be regarded as substitutes for one another in the everyday use of that term, we shall not necessarily find that where the cross elasticity is negative, the goods would be regarded as complementary in everyday usage. First, suppose that good Y is margarine and X butter, and that Y is a Giffen good in the price range OP to OP_1. When the price of margarine falls to OP_1, less margarine and more butter is bought: the demand curve for butter moves to the right, indicating that butter is being substituted for its now relatively cheaper though 'inferior' substitute margarine. The cross elasticity of demand between butter and margarine is therefore negative. Secondly, suppose that the household buys only two goods X and Y and that its planned expenditure remains the same. If the price of Y falls, and if the demand for Y is relatively inelastic, then less will be spent on Y at a lower price and more will be spent on X. That is, the demand curve for X will shift to the right. In both these examples, in the language of our indifference analysis, the positive income effect of the fall in the price of Y on the quantity of X that is demanded outweighs the tendency for Y, which is now relatively cheaper, to be substituted for X. We may, therefore, associate negative cross elasticities of demand not only with complementarity but also with relatively strong income effects.

CONCLUSION

We have described how the purchase plan of a household might be revised when there is any change in the price of one of the goods that it buys, or in its planned consumption expenditure, or in its tastes and preferences. We have shown how the *extent* of the revision entailed by any one of these changes can be measured. From the whole series of purchase plans each

appropriate to a different price for some good X (the prices of all goods other than X, tastes and preferences and planned consumption expenditure all remaining the same) we extracted the demand for X. The total demand for X summarises the role that households play in determining the relative price of X as they implement their decisions to buy it. In the next two chapters, we shall study the sales plans of firms, and attempt to discover the part which firms play as they implement their sales plans, in determining the prices of the things that they sell.

APPENDIX TO CHAPTER 1

We justified in part the choice of indifference curves that were convex to the origin on the grounds that only with indifference curves of this shape can we illustrate the typical reactions of household to changes in their incomes or in the relative prices of the things that they buy. We must now show explicitly that indifference curves that are either concave or straight lines when viewed from the origin would suggest reactions that are contrary to those which we observe.

In Diagram 19, we have an indifference map in which each of the curves is of the shape I_2 in Diagram 3. Let us now try to illustrate the purchase plan of the household on this diagram as

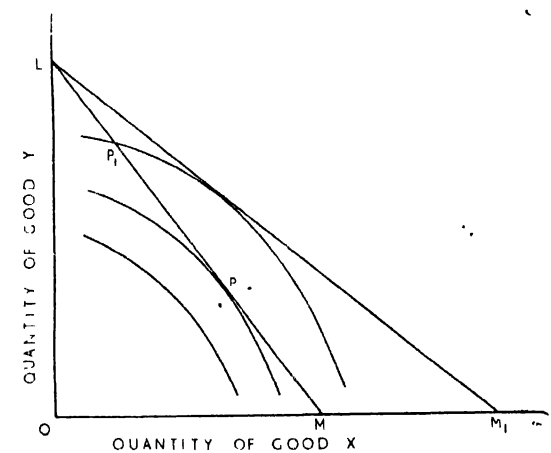


DIAGRAM 19

we did in Diagram 9. If we know the household's planned consumption expenditure and the prices of X and Y, we can draw the budget line LM. This touches an indifference curve at P. But the household will clearly not choose the plan P, for the further one moves along the budget line in either direction from P, the higher are the order of the indifference curves that are reached. P_1, for example, lies on a higher indifference curve than P, and therefore represents a combination of X and Y that the household will prefer to P. In these circumstances, the household will plan to buy only one good: it will buy either X or Y depending on whether L or M lies on the higher indifference curve. If L lies on the higher curve, and if the price of X falls, then the household may stop buying Y and buy only X, for M_1 is likely to lie on a higher indifference curve than L.

If the indifference curves are straight lines when viewed from the origin, we get the same kind of results. If the slope of the budget line is less than the slope of the indifference curves, then the household will buy only X; if the budget line is steeper than the indifference curves, the household will buy only good Y. If there is a relative change in the prices of the goods that the household might buy, then the household might switch its purchases completely from one good to the other.

We know that flesh-and-blood households do not generally buy only one good, nor would they usually cease buying that good altogether and concentrate wholly on some other good, if its relative price were to rise. It is for this reason (*inter alia*) that we reject indifference curves of the shapes I_2 and I_3 in Diagram 3, and assume that all indifference curves are convex to the origin.

CHAPTER 2

The Sales Plan of the Firm: Short-run

The economic unit for the sale of goods and services is the firm. The firm may be a small farmer, cultivating his own land with the help of the members of his family or a village barber plying his shears in his own front parlour. It may be an industrial, commercial, financial or agricultural undertaking, owned by one person, by a partnership of two or more persons, by a private company, by a public and limited liability company; or it may be nationalised and so owned by all of us and operated on our behalf either directly or indirectly by a local, or by the central, political authority. In economic analysis, we are primarily interested in what happens when firms revise their sales plans and their purchase plans. We are not directly interested in whether the firm is owned by one or by many, or in whether it is large or small, strong or weak, adolescent or moribund. These qualities demand our attention only in so far as they affect the sales and purchase plans which the firm chooses — that is, in so far as they influence the objective which the firm is pursuing.

All firms, irrespective of their size, organisation and objectives, must make sales and purchase plans. In the sales plan of the going firm are set out the quantities of each of its products that it plans to sell during any period of time, and the price per unit at which it hopes or plans to sell each of them. A typical arable farmer, for example, may make the following sales plan for the season lying ahead: 30 tons of potatoes which he hopes to sell at £10 per ton; 10 tons of cabbages at £15 per ton; 15 tons of turnips at £18 per ton; 35 tons of sugar beet at £20 per ton, and 250 bushels of corn which he hopes to sell at £3 per bushel. The total value of his planned sales adds up to £2170, which is his expected gross revenue for the ensuing season. If

we use 'p' to represent the expected selling price, 'q' to represent the quantity that he hopes to sell, and the subscripts 1, 2, 3, ... n, to designate particular products, then the sales plan of the firm will have the following general form:

$$p_1 \cdot q_1 + p_2 \cdot q_2 + p_3 \cdot q_3 + \dots p_n \cdot q_n = R,$$

where R represents the gross revenue that the firm hopes to get from implementing the plan as a whole.

If this plan is to be realised, the farmer must buy seed and fertiliser and hire labour, equipment and land, or use his own if he has them. The quantities of each of the things that he must hire or buy and the price which he expects to have to pay for each of them are set out in the purchase plan. If we use 'f' to represent price, 'x' to represent quantity to be bought, and the subscripts 1, 2, 3, ... m, to designate particular things that the farmer must buy or hire, then the general form of his purchase plan will be as follows:

$$f_1 \cdot x_1 + f_2 \cdot x_2 + f_3 \cdot x_3 + \dots f_m \cdot x_m = C,$$

where C represents the total expenditure that he feels he must incur if his sales plan is to be implemented.

For brevity's sake, we shall call C his total costs, and to avoid the ascientific flavour of the phrase 'the things that the firm buys or hires' we shall call these things factors of production, or productive services or inputs. We shall also hereafter refer to the 'things that firms sell as products or outputs. We can never uniquely identify any particular service or commodity as either a product or a factor, for one firm's product is frequently another firm's factor. The product of a coal-mine, for example, is a factor or input to firms making steel; the products of steel mills are factors of production to firms making motor-cars, omnibuses, steel rails and refrigerators, and the products of these are, in their turn, factors of production to transport undertakings and the makers of ice-cream. For our purpose, a good or service will be treated as a factor of production to any firm whose purchase plan contains it, and as a product for any firm that plans to sell it.

In describing the sales and purchase plans of a firm, we must take care to include not merely those products which the firm plans to sell to, or those inputs which it plans to buy from, per-

sons other than its owners. If the members of the farmer's household plan to eat 10 cwts. of potatoes off the farm during the season, then these 10 cwts. must be included in his sales plan. If a farmer uses his own land, its services must be included in his purchase plan. If the owner of a shop is also its manager, we must view him as buying managerial services from himself. If a firm owns machines, then their services must be included in the purchase plan. In the sales plan, we include all the products that the firm produces so that the plan is independent of the manner in which the products are disposed of. In the purchase plan are included all the factors of production that are used in producing the outputs that are listed in the sales plan.

Now, there are many sales plans that a firm might make for any period of time. The farmer's sales plan that we described above was only one of many plans that the farmer might have laid for the period of time in question. In an alternative sales plan, he might have planned to produce and sell more potatoes and fewer cabbages and turnips, more turnips and less corn, or more sugar beet and less potatoes and cabbages. And for each sales plan that the farmer might have chosen, there were available many alternative purchase plans. The sales plan described on pages 44–5 might have been put into effect with much labour and little land and equipment, or with much equipment and little land and labour, or with any one of an almost infinite number of combinations of these three kinds of factor of production.

The sales and purchase plans that a firm will actually choose from the many alternative plans that it might make, will depend on the objective that the firm is pursuing and on the length of time to which its plans are related. The goals of firms are influenced by the *mores*, laws and customs of the societies in which they operate. In any given social context, the objective that a firm will pursue will depend on the personality of those who own and control it, on the organisation of the firm, and on the structure of the markets in which it sells its products and buys its inputs. If it is only one of a very large number of firms selling a particular product, then to survive it may have to choose sales and purchase plans that promise the maximum excess of gross revenue over total costs. If it is the sole seller of a product, it may choose plans that promise only a 'satisfactory' or 'reason-

able'. return. The absence of other sellers of the same product may allow its owners to indulge their individual idiosyncracies: they may prefer a quieter life with the promise of rather smaller rewards to a more energetic and exacting existence and more money. A public enterprise, like British Railways, may be required by law to so organise its operations that its total revenues just cover its total costs.

If we are given the objective of the firm, the range of sales and purchase plans from which it may choose will depend on the period of time for which it is planning. In general, the shorter the period of time to which the sales and purchase plans are related, the narrower will be the range of choice, and vice versa. The length of the planning period affects the contents of the purchase plan in two ways: it affects the physical quantities of the different inputs that the firm might use and it in part determines the sums of money that the firm must disburse for their use. Thus, if a firm hires its operative labour and buys its raw materials in weekly contracts, and its other factors on contracts with a longer currency, then for periods shorter than one week, the firm cannot reduce the quantities of inputs at its disposal. It might be able to buy more of some of them, but even here the limits are narrow, for it may take time to find suitable labour to hire and to seek out new sources of more raw materials. The firm need not, of course, use all the inputs at its disposal — but even if it uses none of them, its costs will be the same, for while the contracts run, the labour-service, etc., must be paid. For such very short periods of time, therefore, all the firm's costs might be *fixed* costs. If the planning period is longer — say, one month — then the firm's range of choice will be wider, for during a month the quantities of all inputs that are hired on contracts that have a currency of less than one month can be increased or decreased. Time, however, exerts its influence not only through the possibility of making, modifying or renewing contracts, but also because time is needed in which to produce the new inputs that the firm may require. Thus, it might take twelve months to build a new factory: for planning periods of less than one year, the input 'factory-space' must be taken as constant.

The influence of time on the number and scope of the different plans from which a firm may choose will be explored more fully

in the next chapter. For the moment, we conclude that the range of choice open to the firm varies directly with time: for very short periods, the range of choice may be virtually zero; for very long periods, it may be virtually infinite. While this relationship between time and the number of alternative decisions that a firm might make is a continuous one, it is customary in economics to explore the role of time by taking three discrete periods:* the instantaneous or market period, the short-period (or short-run), and the long-period (or long-run). In the instantaneous period, the sales and purchase plans are data. In the short-run, the quantities of *some* of the inputs that the firm uses can be increased or decreased: it is usually assumed that operative labour and raw materials are variable while the quantities and qualities of the plant, machinery and managerial labour are fixed. In the long-run, the quantities and qualities of all the inputs that the firm might use can be varied. In this chapter, we shall concentrate on the alternative sales and purchase plans that might be made for the short period, and in the next chapter on the range of choice open to the firm in the long-run.

We have now described the sales and purchase plans of the firm. Our general observations of economic activity confirm that all firms have these plans, either explicitly or implicitly, for all firms buy and sell. We assume that the sales plan and the purchase plan which a firm decides to implement are chosen because they promise to fulfil the objectives that the firm is pursuing. As economists, we are not generally concerned with the precise contents of a firm's plans at any particular time. We are interested in how these plans are revised when the data on which they are based alter and in how these revisions help to determine the prices of the firm's products and factors. In this chapter, we shall describe the data on which the sales plan of a firm is based, and we shall show how the sales plan is revised when any datum changes.

The data on which a firm's sales plan is based are partly technical and partly economic. The technical data are of two kinds. First, there are the different *methods* by which any commodity may be produced. Thus, woollen cloth may be woven

* The dangers inherent in arbitrarily dividing a continuous process into discrete intervals are perhaps best described (though in a different context) in L. Tolstoy, *War and Peace*, Volume III, Chapter 1, pages 3–4. (Translated by L. and A. Maude, Oxford University Press.)

on hand-looms in a cottage or on power looms in a large factory. Each of these is a separate method of producing woollen cloth. We generally find that both the quality of the product and the kinds of productive services used to produce it vary from one method or technique of production to another. Cloth woven on a hand-loom is recognisably different from cloth woven in a factory and is rated differently by consumers. The labour skills and the kinds of equipment and machines required for factory production differ in quality from those needed in handicraft production. Second, with each method of production varying quantities of woollen cloth can be produced by varying the quantities of the appropriate productive services that are used. Thus, more cloth could be woven on a hand-loom if more yarn and more hours of labour-service were employed: a factory could increase its output of cloth, if more operatives were hired, more yarn bought and more looms brought into use. These technical data — the knowledge of the different methods by which a commodity might be produced and of how the quantity that is produced will behave, within each method, as the quantities that are used of the appropriate productive services are varied — constitute the production possibilities open to the going or prospective firm. The technical data are engineering relationships. As economists, we are not primarily concerned with why these are what they are or with why they change through time: our main concern is to show how these data form the basis of the sales and purchase plans of firms.

From the whole range of production possibilities, the firm chooses what products to produce, the method by which to produce them and what quantities of them to offer for sale, in the light of the economic data. The economic data are the prices at which the firm expects to be able to sell the different commodities that it might produce, the prices which it expects to pay for the different productive services it would require and the objective that the firm is pursuing.

It is worth noticing, *en passant*, that the data on which a household bases its purchase plan — and which we described in Chapter 1 — may be classified in a similar way. The equivalent of the technical data for the household is its awareness of the satisfactions it expects to derive from each possible combination of the commodities and services available for its consumption.

From the whole range of consumption possibilities, the household chooses what goods to buy, and in what quantities to buy them, in the light of the economic data. The economic data for the household are the prices of the goods available for its consumption, the household's planned consumption expenditure and the objective that it is pursuing. In Chapter 1 we portrayed the pattern inherent in the consumption possibilities by means of an indifference map, and illustrated the choice by the household of that combination of goods which promised to leave it as well-off as it could be in the light of the limitations imposed by the given sum of money available for expenditure and by the fact that a price had to be paid for each good that it might buy.

In our attempt to discover how a firm will revise its sales plan if there is any change in the planning data, we shall proceed by roughly the same steps as in Chapter 1. First, with the help of a simple example, we shall illustrate graphically the patterns that have been observed in the production possibilities which are open to a firm that is already in existence. Second, we shall introduce the economic data and show how the prices of the factors of production determine the quantity of each of these factors that the firm would use to produce each possible output, and how the selling price of the output determines which of the possible outputs the firm will produce and sell. Third, we shall then be in a position to show how the sales plan will be revised when any economic or technical datum changes. In this chapter, our analysis will be confined to the short-run — that is, to a period during which the quantities of only some of the firm's inputs can be varied.

THE PRODUCTION POSSIBILITIES OPEN TO THE GOING FIRM

For our initial example, we shall suppose that the firm has decided to produce only one product, namely, a particular quality of woollen cloth, and that it has already chosen the method by which to produce it. In the next chapter, we shall illustrate how these decisions are made. Having implemented these decisions, the firm has a factory building of a given size and design, and in it have been installed a certain number of machines and a given quantity of other equipment. To produce its product, we shall suppose that certain qualities of two other

inputs, *X* and *Y*, must be employed, and that the quantity of the product may be varied by varying the quantities of these that are used. We may think of *X* and *Y* as being two different kinds of labour-service, or of *X* as being labour-service and *Y* yarn.

In Diagram 20, we draw two straight lines at right angles to one another. We measure quantities of *X* that might be used per week on the horizontal axis, and the quantities of *Y* per week on the vertical axis. Any point that lies between these axes represents a combination of a particular quantity of *X* with a particular quantity of *Y*. The planning unit or planner of the firm — it or he is usually called the entrepreneur — will expect each combination of *X* and *Y*, when used in conjunction with the firm's buildings, machines, equipment and executive and managerial labour, to produce a certain quantity of cloth. Beside each point that lies between these axes, we can write the amount of cloth that the quantities of *X* and *Y* that it denotes would be expected to produce. Thus, the combination of the two factors denoted by *P* might promise an output of 4,000 yards per week; the combination *Q* might promise an output of 6,000 yards per week and that denoted by *R*, 3,000 yards per week. When this is done, we have a visual representation of the production possibilities open to the firm in our example. From all these possible outputs the firm must choose one, namely, that

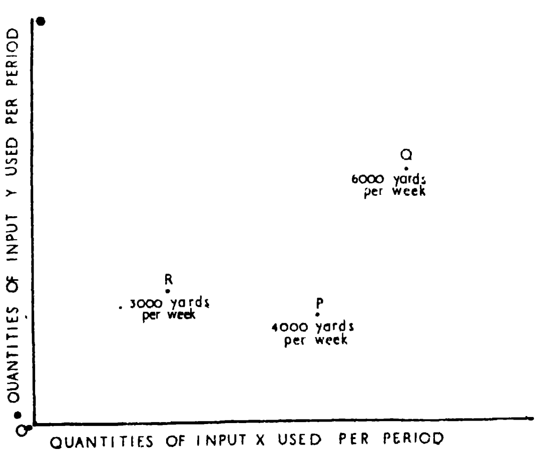


DIAGRAM 20

quantity of cloth and the quantities of X and Y necessary to produce it, which if sold and bought respectively, would promise to achieve the firm's objective.

We shall be helped in illustrating the choice on our diagram if we first order the different points that lie between the axes in terms of the quantity of cloth they represent. We can do this in the same way as, in Chapter 1, we introduced some semblance of order into the infinite number of points each representing a combination of two goods possessing a particular degree of attractiveness to the household: that is, by joining together all combinations of X and Y that promise to yield the same quantity of cloth. These lines are called *isoquants*, or *product-indifference curves*, or *iso-product curves*.

The shape of the isoquant curves may be established either directly by experiment, or indirectly in the same way as we established the shape of indifference curves. In Chapter 1, we deduced the shape of indifference curves and we derived their properties from introspection and from the observed reactions of households to changes in the data on which their purchase plans were based. Direct estimation was impossible, for the 'outputs' of psychic satisfactions, into which the household transforms the 'inputs' of commodities and services, could not be cardinally measured. The outputs of a firm, however, can be measured in cardinal units, and this enables us to discover the shape and properties of the iso-product curves directly. In this chapter, we shall use the direct method.

A production engineer could estimate, or could establish by a series of controlled experiments, what quantities of cloth could be produced by the different possible combinations of X and Y, when used with the firm's buildings, machines, equipment and other productive services. The output expected from any given combination of the inputs X and Y will, of course, depend on the way in which it is combined with the other inputs available to the firm. Thus, if X is labour-service and Y yarn, the estimated output from using 25 hours of labour-service and 100 pounds of yarn on one of the firm's machines will be different from what it would have been had they been set to work on three machines. Indeed, each particular combination of X and Y could produce a whole range of different outputs depending upon how it is combined with the other, fixed inputs. To avoid this indeter-

minancy, we shall assume that the production engineer in making his calculations will always combine any given combination of X and Y with the other, fixed inputs in such a way as to promise the maximum output. The results of these experiments could be plotted on a diagram such as Diagram 20, in much the same way as a cartographer plots the altitude of particular points on a map. When this is done, and when all combinations of X and Y that promise to yield the same output of cloth per period are joined together, the resulting iso-product curves will generally have the shape and properties illustrated in Diagram 21.

First, their shape. Each iso-product curve will be elliptical.* Curve 1 passes through all the combinations of X and Y that will yield 1,000 yards of cloth per week when used in conjunction with the firm's other productive services. If the firm wishes to produce 1,000 yards per week, it will restrict its choice of the quantities of X and Y to use to the combinations of these services represented by points on the arc AB, for if it chose to produce 1,000 yards per week of cloth by any combination other than one of these, it would be using either more of X or more of Y or more of both than was necessary. For example, LZ units of Y and MZ units of X could produce 1,000 yards of cloth per week, but the same output can be obtained with the same quantity of X by using only LR units of Y. Or 1,000 yards of cloth can be woven in a week with either PU of Y and NU of X, or with PU of Y and NS of X. Faced with such alternatives, we shall assume that the entrepreneur will al-

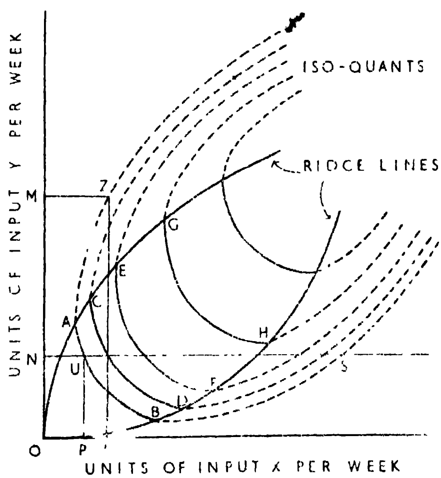


DIAGRAM 21

* Here (and hereafter), the words 'ellipse' and 'elliptical' are not given their strict mathematical meanings. They are used roughly to mean 'egg-shaped' — to draw attention to the fact that each isoquant approaches and then recedes from each axis.

ways choose that combination which involves the least expenditure of productive services: if he wishes to produce 1,000 yards per week, he will always decide to do so with quantities of *X* and *Y* denoted by one of the points on the segment *AB* of the iso-product curve.

This relevant segment *AB* of the iso-product curve is convex to the origin. It has this shape because to some extent the inputs *X* and *Y* can be substituted for one another in the manufacture of cloth. The degree to which these two inputs can be substituted for one another varies with the quantities of them that are already being used. Thus, as we travel north-westwards along the iso-product curve from *B*, fewer and fewer units of *X* can be substituted for a unit of *Y* if the output of cloth is to remain the same at 1,000 yards per week. In Diagram 22, if the firm were making 1,000 yards per week with the quantities of *X* and *Y* denoted by *B*, that output could be maintained by using *GB* less of *X* and an additional unit, *GH*, of *Y*. If it were producing this output with the combination *H*, the output could be maintained by employing *JH* less of *X* and *JK* more of *Y*. And so on along the curve: as the quantity of *X* that is being used is reduced, less and less of *X* can be substituted for each additional unit of *Y*, if the total output is to remain unchanged.

This property of the typical iso-product curve is usually expressed in technical terminology by economists. The reduction

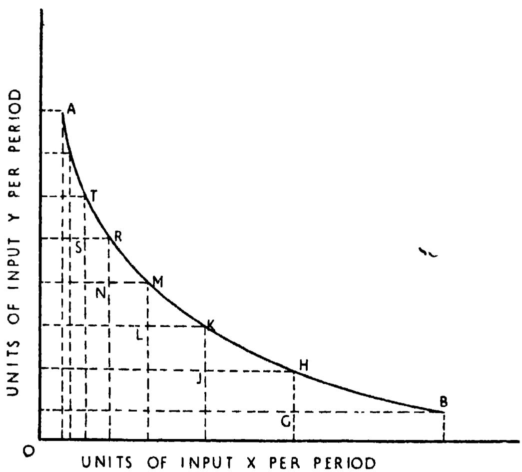


DIAGRAM 22

in the quantity of input X which in the estimation of the firm would just be compensated for by an additional unit of input Y is called the ***marginal rate of technical substitution*** of X for Y. In Diagram 22, in the range BH of the iso-product curve, the marginal rate of technical substitution of X for Y is equal to GB/GH. In the range HK, the marginal rate of technical substitution of X for Y is JH/JK. Our statement that the relevant portion of an iso-product curve is convex to the origin, when translated into this new terminology, becomes a statement that the marginal rate of technical substitution of the input X for the input Y decreases, as Y is continuously substituted for X. At B, the $MRTS$ of X for Y is infinitely large; at A, the $MRTS$ of X for Y is zero.

The fact that each iso-product curve is elliptical is not the result of any economic phenomenon. It illustrates solely the physical, technical relationships that exist, in our example, between inputs of X and Y and the output of cloth. These relationships are data for economics and *quaesita* for other disciplines. While we need not strive to explain the character of these relationships we may, nevertheless, confirm that they are not contrary to what we would expect. In the production of most commodities and services, some inputs can generally be substituted for some others. Thus, in hairdressing, the same number of heads can be shorn by electrically operated clippers and fewer barbers, or by more barbers with manually operated clippers and shears. In most industries, the proportions between operative and supervisory labour can be varied, and the same output produced with relatively more of the former and less of the latter, or with rather more of the latter and less of the former. In our example, if X is operative labour and Y is yarn, then X and Y can be substituted for one another; for with more labour-service the rate of wastage of yarn will be reduced. This is what is happening in the range of substitution near B on the iso-product curve in Diagram 21.

In some industries we find that the proportions in which certain inputs can be combined are fixed and invariable, and that these inputs cannot therefore be substituted for one another in the production of those industries' products. Over the whole field of industrial activity, such circumstances are probably rare, though they may be common in some particular industries

like those producing chemicals. When we find that two inputs must be used in fixed proportions to one another, the isoquants will be parallel to the axes throughout their length, as in Diagram 23. Where two or more inputs must be used in fixed proportions, we may treat them as being a single input; and we usually find that this composite input can to some extent be substituted for some other input. For example, to produce some chemical product, sodium and chlorine may have to be combined in equal proportions, but in the manufacture of that product it will generally be possible to substitute some other input, such as labour-service, for the composite input one unit of which consists of equal quantities of chlorine and sodium.

While substitution is generally possible, it cannot be carried on indefinitely. In our example, if we again suppose that X is labour-service and Y yarn, the two can be substituted for one another, but sooner or later there comes a point beyond which, with the given quantities of other productive services at our firm's disposal, we would expect those who render labour-service to impede one another. If the output is to be maintained, more yarn must be used. At B in Diagram 21, the limit of substitution of labour-service for yarn is reached: beyond B the iso-product curve veers away from the horizontal axis, indicating that more labour-service proves cumbersome and must be offset by more yarn. Similarly, beyond A, yarn is cumbersome and if the pro-

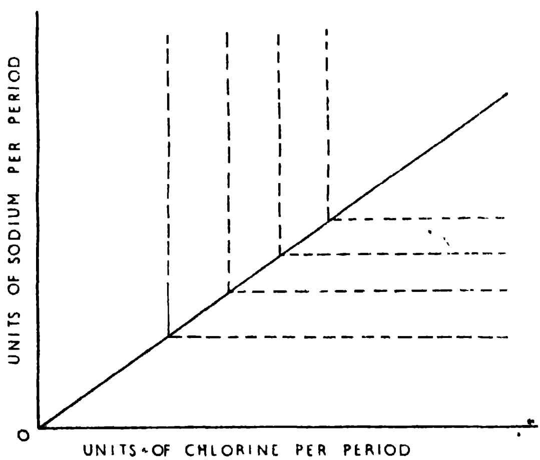


DIAGRAM 23

duct is to be maintained at 1,000 yards per week, more labour-service must be used.

We have seen that the locus of all combinations of X and Y that promise an output of 1,000 yards of cloth per week, when used with the buildings, machines and other factors at the firm's disposal, will be elliptical,* and that if the firm should wish to produce this weekly output it will be restricted in its choice to combinations lying on the segment AB of the isoquant. If the firm is contemplating an output of 2,000 yards per week it will choose to do so with one of the combinations of X and Y that lie on the segment CD of isoquant II; if it is considering a weekly output of 3,000 yards per week, it will decide to use quantities of X and Y denoted by one of the points lying on the segment EF of isoquant III. And so on. When all points such as B, D, and F, and all points such as A, C, and E, are joined together by curves passing through the origin O, we get a visual representation of that part of the whole isoquant map to which the firm will restrict its choice on purely physical or technical grounds. Lines such as $OBDF$ and $OACE$ are called *ridge lines*. It is from the production possibilities enclosed by the ridge lines that the firm will choose what output to produce and what quantities of X and Y to employ in producing it.

This concludes our description of the shape of the iso-product curve: next, we must describe the typical relations that will obtain between one isoquant and another. Let us draw through some point M on the horizontal axis in Diagram 24 a straight line parallel to the vertical axis. As we move northwards along this line we can read off the quantities of cloth that will be produced by increasing quantities of input Y when used with the firm's buildings, equipment and other factors, and the given quantity OM of input X. Thus, when OA units of Y are employed with the given quantities of other productive services 1,000 yards of cloth would be produced per week; if OB units of Y are used, the output will be 2,000 yards per week. The weekly output of cloth would continue to increase until OF units of Y are being used, when it reaches a maximum of 6,000 yards per week. Thereafter, with further increases in the input of Y, the weekly output declines. The firm, through circumstances beyond its control, may find itself restricted to the production

* See footnote, page 53.

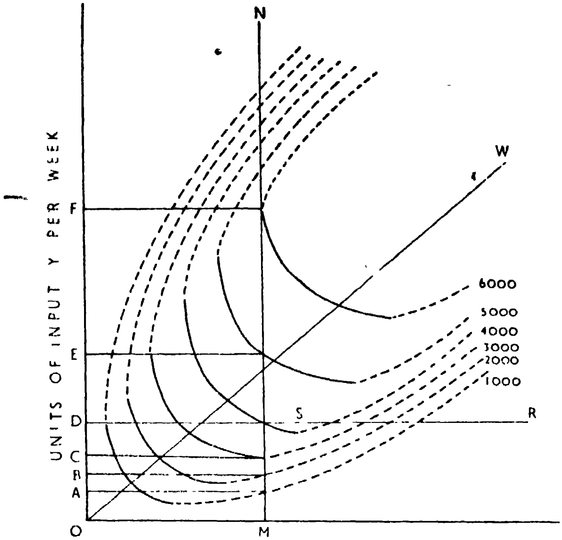


DIAGRAM 24

possibilities lying on the line *MN*: if *X* were labour-service, the firm, through trade union intransigeance, may be forced (for a while at least) to employ *OM* hours of labour-service per week. If this should happen, the firm will employ some quantity of *Y* not greater than *OF* with its quota of labour-service and its other factors.

It will usually be found not only that output increases as we move along some line such as *MN*, but that the rate at which output rises follows a general pattern. This pattern will be clearer if we plot the behaviour of weekly output against the input of *Y* on a separate diagram. This is done in Diagram 25. On the horizontal axis, we measure the quantities of input *Y* used per week in the same units and on the same scale as on the vertical axis of Diagram 24. If we think of the isoquant map in Diagram 24 as representing the configuration of a 'hill' of production, in the same way as contour lines on a map show the configuration of a mountain, then Diagram 25 portrays a cross-section of the 'hill' along the line *MN*. It shows what the 'hill' would look like when viewed from some point due west of *MN*.

The distinguishing characteristics of the behaviour of weekly

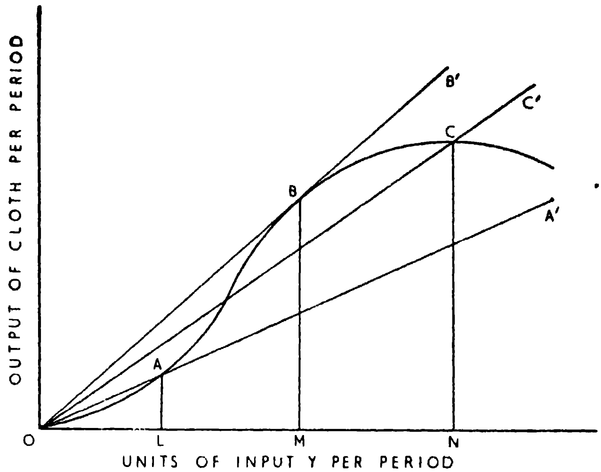


DIAGRAM 25

output become clearer if we compare the rate of increase of weekly output with the rate of increase in the input of *Y*. We can do this visually, and thus avoid laborious arithmetic calculation, in the following manner. At any output, we can discover whether a small change in the input of *Y* promises a greater or smaller proportional change in output by drawing a straight line from the origin to the point on the curve corresponding to that output. If the line cuts the curve from above, then the proportional increase in output is greater than the proportional increase in the input of *Y*; if the line cuts the curve from below, then the proportional increase in input exceeds the proportional increase in output. Thus, in Diagram 25, *OL* of *Y* promises an output of *LA* per period. The straight line *OAA'* shows what the behaviour of output would be if changes in the input of *Y* (from its initial level *OL*) caused equi-proportional changes in output. In the region of *A*, the total product curve rises more steeply than the guide line *OAA'*, showing that changes in the input of *Y* (from *OL*) promise changes in output of a larger proportion. Similarly, *OM* of *Y* promises an output of *MB* per period; at this output, the guide line *OBB'* is a tangent to the total product curve at *B*, showing that small percentage changes in the use of *Y* (from the level *OM*) promise changes of the same percentage in output. Finally, at the output corresponding to *C*, the guide line cuts the curve from below, indicating that

small variations in the input of *Y* from *ON* promise smaller proportional changes in output.

This pattern of behaviour of weekly output, that appears when varying inputs of *Y* are combined with given quantities of *X* and other factors, is summarised in the Law of Non-Proportional Returns. This law states: With a given method of production, the application of further units of any variable input (in our example, *Y*) to a fixed combination of other factors (in our example, *X*, buildings, machines and other equipment) will, until a certain point is reached, yield more than proportional increases in output, and thereafter, less than proportional increases in output. This law is sometimes called the Law of Variable Proportions to emphasise the conditions that must obtain before the non-proportional returns can be expected to appear. The weekly output will only follow the pattern that this law describes where it is possible to use varying quantities of one factor with fixed quantities of other factors. That is, it must be possible to combine the variable factor in varying proportions with the fixed factors.

We would find the same pattern in the behaviour of weekly output if varying quantities of *X* were used with a fixed quantity of *Y* and the other factors at the firm's disposal, or if varying quantities of *X* and *Y* were used with the given buildings, machines and equipment, etc., of the firm. If input *Y* is yarn, the firm through a strict government rationing scheme, for example, which allowed it a weekly quota of only *OD* pounds of yarn, might be restricted to the production possibilities that lie on the line *DR* in Diagram 24. In these circumstances, the firm will employ some quantity of *X* not exceeding *DS* units per week, for as more *X* is used with the fixed quantity of yarn and other factors, weekly output rises to a maximum when *DS* of *X* is being used, and thereafter declines. The rate at which the output rises as more of *X* is used will follow the pattern described by the Law of Non-Proportional Returns. The same pattern would be discernible, and for the same reasons, in the production possibilities lying on a line such as *OW*: as the quantities of *X* and *Y* that are used with the firm's buildings, machines, etc., are increased, after a certain point, the increases in weekly output will be proportionately less than the increases in the inputs of *X* and *Y*.

We have now taken the first step described on page 50 above. A two-fold pattern is discernible in the production possibilities open to the going firm. First, the locus of the different combinations of the two variable inputs, *X* and *Y*, with which a given output of cloth can be produced per week, is an 'ellipse'. Second, the relationship between successive loci or isoquants follows the pattern described by the Law of Non-Proportional Returns. The firm will restrict its choice among the production possibilities, so ordered, to those lying between the ridge lines, for the production of any output by a combination of *X* and *Y* lying within the ridge lines involves the use of less *X* and *Y* than the production of that output with any combination lying beyond the ridge lines. From all the possible weekly outputs, the firm must choose one, namely, that quantity of cloth and the quantities of *X* and *Y* needed to produce it, which, if sold and bought respectively, would promise to fulfil the firm's objective. The next step is to show how, in the light of the prices that have to be paid for *X* and *Y*, the firm chooses the particular combination of the two inputs by which it would plan to produce each possible output, and then to illustrate how the price per yard at which the firm expects to sell its output helps it to decide what particular output to produce and sell.

THE PRICES OF INPUTS AND THE CHOICE OF HOW TO PRODUCE EACH OUTPUT

If we know the prices at which the firm expects to be able to hire or buy the inputs *X* and *Y*, we can show on the same kind of diagram as that on which we drew the isoquant map the different combinations of *X* and *Y* that the firm could buy with various sums of money. On the vertical axis in Diagram 26, we measure units of input *Y* and on the horizontal axis units of *X*. If *X* can be bought for 5s. per unit, then for a sum of, say, £10, 40 units of *X* can be bought per week. This is represented by *OL*. If *Y* can be bought for 2s. per unit, then for the same sum 100 units of *Y* can be bought. This is shown by *OM* on the diagram. The straight line joining the points *L* and *M* will pass through all combinations of *X* and *Y* that the firm can buy with £10, if it spends all that sum of money on their purchase at the given expected prices. An infinite number of such lines can be drawn, all parallel to one another, and each representing the various

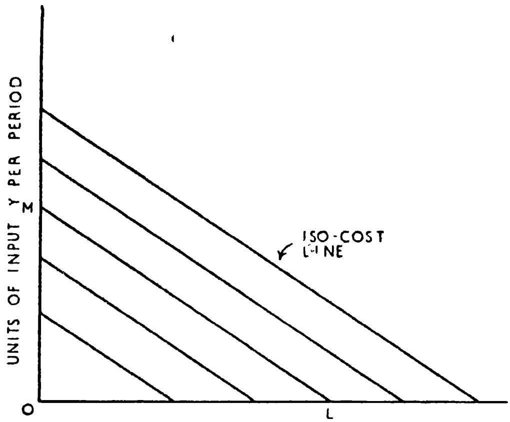


DIAGRAM 26

quantities of the two inputs that can be acquired for a particular sum of money. These lines are called iso-cost lines.

In Diagram 27, some of these iso-cost lines are drawn between the same axes as the firm's isoquant map. Each isoquant shows us the different combinations of X and Y with which a particular weekly output might be produced. The price or iso-cost lines show us what each combination of X and Y would cost the firm. We shall assume that the firm will always plan to pro-

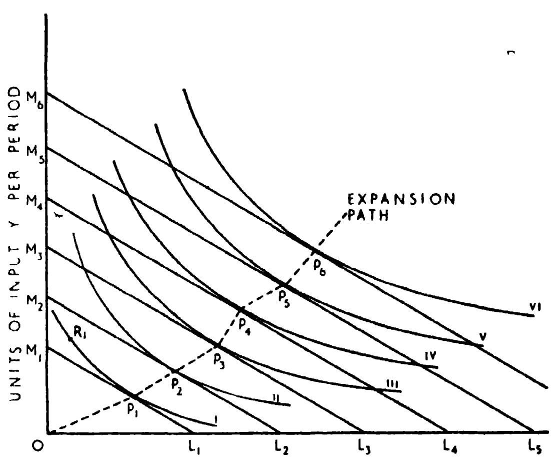


DIAGRAM 27

duce each output with the quantities of the inputs X and Y that cost least. Thus, if the firm wishes to produce 1,000 yards of cloth per week, it will plan to do so with the quantities of the two inputs denoted by P_1, for any other combination by which 1,000 yards might be produced will lie on a higher iso-cost line and will therefore cost more. The combination R_1, for example, would produce 1,000 yards per week, but R_1 lies beyond L_1M_1 and would cost the firm more than P_1. As we move along the isoquant I in either direction from P_1, the quantities of X and Y that must be bought to produce 1,000 per week become progressively more expensive. Similarly, if the firm wishes to produce 2,000 yards per week, it will plan to do so with the quantities of X and Y denoted by P_2; if it were planning to produce 3,000 yards per week, it would do so with the quantities of the two inputs denoted by P_3, and so on.

A property of the purchase plan by which the firm would plan to produce each weekly output is that the marginal rate of technical substitution of Y for X is equal to the ratio of the price of X to the price of Y. In Diagram 27, the $MRTS$ of Y for X when the firm is planning to buy the quantities of these two inputs denoted by P_1 is equal to the slope of the isoquant at P_1, which is the same as the slope of the price line L_1M_1. The slope of L_1M_1 is equal to OM_1/OL_1. We know, however, that OM_1 equals a certain sum of money, say £10, divided by the price per unit of Y, and that OL_1 equals the same sum divided by the price per unit of X. The slope of L_1M_1 is then equal to:

$$\frac{£10}{P_y} \bigg/ \frac{£10}{P_x} = \frac{P_x}{P_y}.$$

To say that for each output the firm chooses the purchase plan which equates the $MRTS$ of Y for X with the ratio of the price of X to the price of Y is merely to say in a different way that it chooses that which promises to minimise its total expenditure on the quantities of X and Y needed to produce that output.

We have now taken the first part of the second step described on page 50 above. From all the possible combinations of input X and input Y with which each output might be produced, the firm will exclude those that lie beyond the ridge lines. From all the remaining combinations of the two inputs with which each output might be produced, the firm will choose that which

minimises its expenditure on them. The line joining the minimum cost combinations denoted by the points P_1, P_2, P_3, etc., is called the *expansion path*, because it shows how the purchases of *X* and *Y* will be increased as weekly output is expanded.

The relationship between weekly output and the firm's expenditure on *X* and *Y*, which is implicit in the expansion path, is shown explicitly in Diagram 28. On the horizontal axis we measure the planned weekly output of cloth and on the vertical axis we measure the expected weekly expenditure on buying the inputs *X* and *Y*. When the expenditure on *X* and *Y* is plotted for each possible weekly output and the points joined together we have the firm's variable cost curve. It is called the variable cost curve, for it shows us how the firm's expenditure on the two variable inputs behaves as output increases.

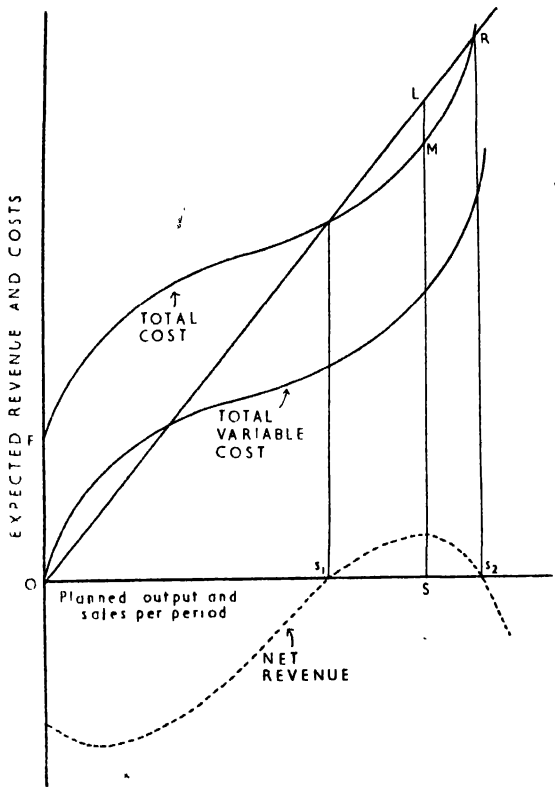


DIAGRAM 28

The peculiarities of the shape of this curve become clearer if we compare the rate of increase in weekly output with the rate of increase in variable costs. This can be done by the method described on pages 58–9 above. At any output we can discover whether a small change in output causes a greater or smaller proportional change in variable cost by drawing a straight line from the origin to the point on the curve corresponding to that output. If this line cuts the curve from below, then a change in output will cause a less than proportional change in costs; if this line cuts the curve from above, then the proportional change in variable costs will exceed the proportional change in output. It is clear from the shape of the variable cost curve in Diagram 28 that for a while increases in output entail less than proportional increases in variable costs, but that after a certain output (which will be between Os_1 and OS) further increases in output involve more than proportional increases in variable costs.

We have not yet established, however, why variable costs behave in this way. The explanation may seem to lie in the Law of Non-Proportional Returns, but this is wholly true only in certain limiting cases. If labour-service, for example, were the only variable factor of production, then, as was shown above, weekly output would at first increase more than proportionately, and then less than proportionately, as the input of labour-service was increased. If labour service could be hired at a given wage-rate, it would necessarily follow that variable costs would at first rise less than proportionately, and then more than proportionately, as weekly output was expanded. This is illustrated in Diagram 28(*a*). The curve shows the behaviour of weekly output as the input of labour-service is increased. If the wage-rate is given, then for hours of labour-service on the horizontal axis we may substitute the *value* of the hours of labour-service. But in our example the value of the labour-service used is the same as variable costs of production. If the diagram is rotated anti-clockwise through 90 degrees, the curve is the mirror image of the variable cost curve in Diagram 28. It is clear, then, that in the limiting case where labour-service is the only variable factor, the shape that we have given to the variable cost curve is wholly explained by the Law of Non-Proportional Returns.

The same argument may be used if the expansion path in Diagram 27 is a straight line passing through the origin. On

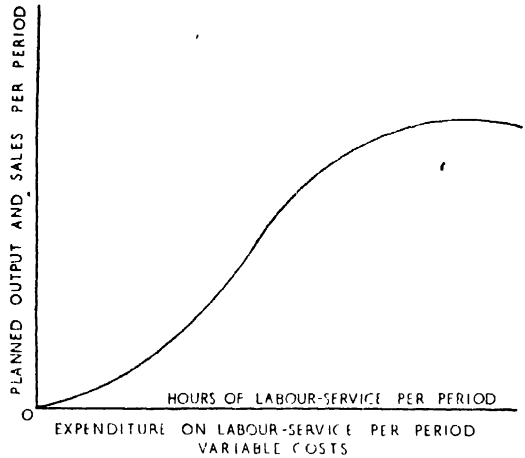


DIAGRAM 28 (*a*)

page 60 above, we saw that if the inputs *X* and *Y* are both increased by the same proportion, weekly output will at first rise more than proportionately, and then less than proportionately. If the expansion path is a straight line passing through the origin, then variable costs will always increase by the same proportion as the variable inputs, when the prices of the variable inputs are given. In this case, therefore, on a diagram similar to Diagram 28(*a*), units of the variable inputs *X* and *Y* on the horizontal axis can be translated directly into units of variable cost: the shape of the variable cost curve will again be wholly explained by the Law of Non-Proportional Returns.

It will generally be found, however, that the proportions in which the variable inputs are combined with one another will change as output is expanded. In this case, there is no way of measuring the percentage increase in the variable inputs needed to effect a given increase in weekly output. If the least expensive way of increasing output by 10 per cent is to hire 12 per cent more labour-service and to buy 8 per cent more yarn, and if there are no meaningful units to which both labour-service and yarn can be reduced, then we cannot say by how much the total variable input has risen. It will be remembered that our argument in the previous paragraphs depended wholly on our ability to do so, for only then was it possible for us to translate changes in variable input into changes in variable costs.

In these circumstances, the shape of the variable cost curve cannot be attributed in the same simple way to the operation of the Law of Non-Proportional Returns. The general tendency described by this law will still be at work, for the variable inputs are being combined with fixed quantities of the firm's other factors, namely, its plant, machinery and equipment, executive and managerial labour. This tendency may be strengthened or weakened by changes in the proportions of the variable inputs to one another. Suppose we are at P_3 in Diagram 27 where a 10 per cent increase in the inputs of both X and Y promise an increase of less than 10 per cent in weekly output. The diagram shows that the cheapest way to increase output by this percentage is by using relatively more of Y than of X. If, however, the unit price of Y is lower than that of X, it is possible that the output on the combination of X and Y denoted by P_4 will be less than 10 per cent greater than that denoted by P_3. Here, the change in the proportion of the variable inputs to each other works against the tendency for the returns to both factors to diminish, so that variable costs rise by a smaller percentage than output. It is probable, however, that the latter tendency soon becomes dominant, and this conclusion is supported both by experience and observation. Statements by businessmen suggest that as output is expanded variable costs rise faster than output, and these statements are confirmed by our observations that higher prices have to be offered for the things that they sell to induce them to plan to increase their outputs.

The variable factors — X and Y — are not the only productive factors that the firm in our example uses in weaving cloth: consequently, the variable costs are not the total weekly expenses that the firm incurs. The inputs X and Y are used with fixed quantities of other resources, namely the firm's buildings, its machines and equipment, its managerial and office staffs. The quantities of these resources at the firm's disposal are the result of a previous decision made by the firm. At some time in the past, the firm decided what quantities and what qualities of each of these inputs were required. The buildings may have been leased, and the machines and equipment hired. If so, the contracts entered into with their owners will probably run for many of our planning periods. Alternatively, these factors of production may have been bought outright by the firm, with monies sub-

scribed by the firm's owners or borrowed by them from others. If the latter, the firm must pay interest to those who have made loans and make provision for the repayment of the principal. If the former, there will be no contractual obligation to pay fixed sums of money to the firm's ordinary shareholders; but the firm must set aside such sums in each period as will enable it to replace these factors when they have worn out or become obsolescent. The contracts by which the firm hires its managerial and executive labour are likely to have a currency of several of our planning periods. We shall call these expenditures that we have listed the *fixed* costs, for their size does not depend on, or vary with, the size of the firm's weekly output of cloth. We shall suppose that these fixed costs run at the rate of £*OF* per week. When £*OF* are added to the variable costs of producing each output on Diagram 28 and the points plotted and joined, we have the total cost curve. This will not be parallel to the total variable cost curve, but will be distant from it at each output by £*OF*.*

THE CHOICE OF A SALES PLAN

The total cost curve in the diagram shows us the minimum sum of money that the firm must disburse each week to produce each possible output. The next step is to show how in the light of its objective and of its expectations of the price per yard at which it hopes to sell its cloth, it chooses what quantity of cloth to produce and sell.

We shall assume that the firm expects to be able to sell any output of cloth it might produce at a given price — that is, that the price of cloth lies beyond the firm's control. In these circumstances, the alternative revenues that the firm might earn will all lie on the straight line *OR* in Diagram 28, for when the price per yard is given, revenue will always increase in the same proportion as sales. From Diagram 28, we can read off both the total costs the firm would incur in producing each output and the revenue it would earn from selling it. We shall assume that the firm will choose that output which promises the maximum excess of expected revenue over expected costs: that is, the maximum net revenue. There is no direct and unambiguous

* While the total and variable cost curves will not be parallel, the tangents drawn to them at any given output will be parallel to one another, for at any output the rate of change in total costs must be the same as the rate of change in variable costs.

evidence to support this assumption: our main justification for making it is that the revisions of sales plans that it suggests concur with our observations of what actually happens when firms revise their sales plans in response to changes in the planning data, and that it is consistent with what we know of the motives of businessmen.

In Diagram 28, the expected net revenue to be earned by producing and selling any output is equal to the vertical distance between the revenue line and the total cost curve. The net revenue from each output is plotted on the diagram. At weekly outputs of less than Os_1, net revenue is negative, for in this range of output the revenue line lies below the total cost curve. As output is expanded beyond Os_1, expected net revenue rises, reaching a maximum when OS is being produced and sold; thereafter, it declines again, becoming negative at outputs greater than Os_2. Since we have assumed that the firm wishes to produce and sell the output which promises to earn for it the maximum net revenue, it will choose an output of OS per week.

We have now completed the second step described on page 50 above. From the production possibilities open to it, the firm will choose the output which, at the given prices for its product and factors, promises to yield the maximum net revenue. The sales plan of the firm is then:

$$OS \text{ yards} \times \text{the price per yard} = R.$$

The purchase plan consists of the quantity of each of the 'fixed' factors at the firm's disposal multiplied by the price the firm has contracted to pay for it, or decided to pay for it, plus the quantities of X and Y that the firm plans to buy multiplied by their respective prices: the sum of these items equals the total costs of producing and selling the output OS. When the firm is implementing these plans, we say that it is in *equilibrium*; in doing so it is planning to earn the maximum net revenue that the circumstances in which it finds itself permit.

So far in this chapter, we have done little more than illustrate diagrammatically the sales plan of the firm. In doing so, we have presented the data on which the sales plan is based in such a way that we can deduce how the sales plan will be revised if any datum changes. It should now be clear that the sales plan will be revised if there is any change in the constellation of pro-

duction possibilities, or in the prices of one or other of the variable inputs, or in the expected selling price of the product, or in the objective of the firm, or in any two or more of these. In the remainder of this chapter, we shall examine how the sales plan will be revised if there is a change in the expected selling price of the product. Changes in the production possibilities and their effects will be examined more fully in the next chapter, and the effects of changes in the price of one or more of the firm's inputs will be postponed until Chapter 5, when we shall be dealing more fully with the firm's purchase plan.

HOW THE SALES PLAN IS REVISED WHEN THE EXPECTED SELLING PRICE OF THE PRODUCT CHANGES

We see from Diagram 28 that the firm in our example, with its given expectations about the price of its product and the prices of its inputs, will decide to sell OS yards of cloth per week. If for any reason there is a change in the entrepreneur's expectations about the price at which he expects to be able to sell his product, then the sales plan will be revised. If the expected selling price is p_1, then the alternative revenues that the firm might earn will all lie on OR_1 in Diagram 29; given the total costs of production, the firm will plan to sell Os_1 per period. If a price of p_2 per yard is expected to obtain, then OR_2 will be the revenue line, and if its costs remain the same, the firm will plan to produce and sell Os_2 per week, for at the new price this promises the maximum net revenue. By drawing in the appropriate revenue line, we can discover the sales plan that the firm would choose at each expected selling price for its product. It is clear from the diagram that planned output will be the higher the higher is the expected selling price of the product, and that planned output and sales will be the lower the lower is the expected selling price.

At a price of p_3 per unit, the revenue line OR_3 touches the total cost curve at A; in these circumstances, the maximum net revenue that the firm can hope to earn is zero, and it can earn that by producing and selling Os_3 per week. By selling Os_3 per week, the firm will just be earning enough to remunerate all the factors of production that it employs. It may be worth the firm's while, indeed, to continue producing and selling even if the expected net revenue were negative. Thus, if the price were p_4, net

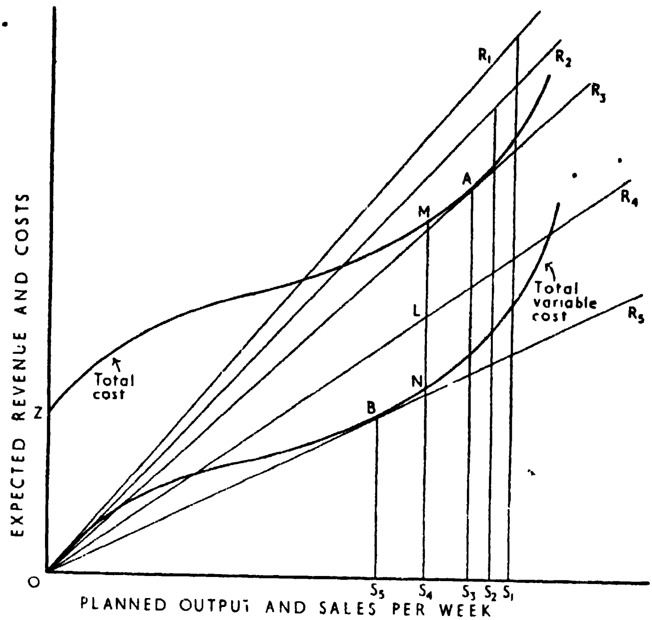


DIAGRAM 29

revenue will be negative at all possible outputs. If the firm decides to produce and sell the output Os_4 where the negative net revenue is least, its total costs will exceed its total revenue by LM per week: the firm must find this sum each week from sources other than its revenue. If the firm did not produce at all, however, it must still honour the contracts made with its fixed factors: when not producing it would have to pay its fixed factors NM per week. Since LM is less than the fixed costs NM, the firm will prefer a weekly production of Os_4 to closure. If the price were to fall to p_5 per unit, at which the revenue line just touches the variable cost curve at B, the maximum net revenue will be earned by producing Os_5 per week. The sale of Os_5 per week yields a revenue just sufficient to cover the variable costs. At this output, the firm will be indifferent as between continued operation and temporary closure. If it closes down temporarily, the firm must find its fixed costs each week; if it produces and sells Os_5 per week at a price of p_5 per unit, its revenue will fall short of its total costs by the same sum.

At any price exceeding p_3 per unit, our firm could continue in production indefinitely, for the revenue earned will always exceed the total costs. At prices less than p_3 and greater than p_5 per unit, the firm can only continue temporarily, for sooner or later its reserves, or the other sources from which it is financing the whole or part of its fixed costs, will be exhausted. At such prices the firm will attempt to revise decisions it has already made: at the given prices for factors it may succeed in lowering its costs of production and so in earning a positive net revenue, by adopting a new method of production, or by altering the scale of its operations within the method which it is already using. The nature of the alternatives open to it will be examined in the next chapter; for the present it suffices to note that the ease with which it may grasp one of these alternatives and the time it will take to do so depend on a host of institutional, contractual, financial and other considerations. Given these considerations, and while the firm's existing contracts with its fixed factors remain in force, the firm will continue to produce and sell at prices greater than p_5 and less than p_3, for the reasons given in the previous paragraph.

This relationship between planned sales and the expected selling price of the product that we have deduced from Diagram 29 is called the firm's weekly *supply* of cloth. In Diagram 30,

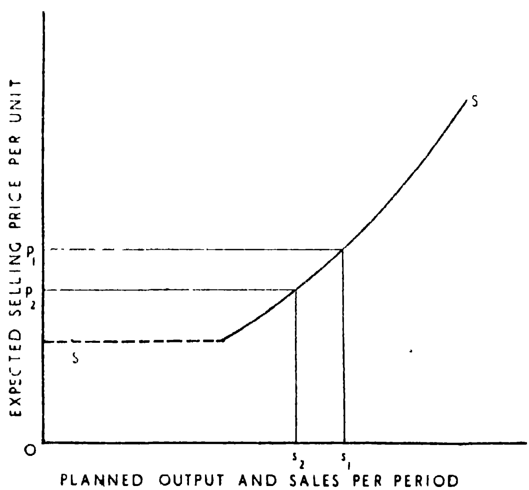


DIAGRAM 30

we measure the expected selling price of cloth on the vertical axis, and the planned sales of cloth per week on the horizontal axis; when our deductions from Diagram 29 are plotted between these axes and the points joined together, we have the firm's weekly supply curve of cloth. This supply curve shows us the sales plan that the firm would choose at each expected selling price for cloth, with its given production possibilities, objective and contractual obligations, and at the given prices it expects to have to pay for its variable inputs. It shows us how the sales plan will be revised if the only datum which changes is the expected selling price of the product.

So far in our diagrams we have illustrated the relationships between weekly output and total costs and total revenue. In economic textbooks, however, it is perhaps more common to portray the relationships between output and average cost and average revenue. These are easily deducible from Diagram 29, and are shown in Diagram 31. On the vertical axis of this diagram we measure sums of money, and on the horizontal axis alternative outputs of cloth per week. The average variable cost of producing any output is equal to the total variable cost divided by the number of units of output: thus, the average variable cost of producing each unit of the output Os_4 in Diagram 29 is s_4N/Os_4 — that is, the slope of the straight line ON.

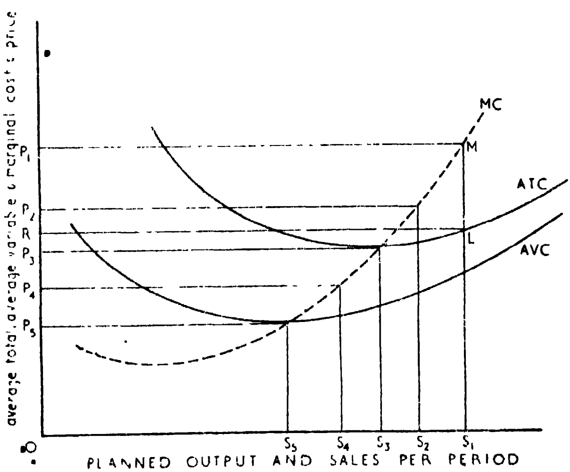


DIAGRAM 31

When this calculation is made for each output and the results plotted and joined in Diagram 31, we have the average variable cost curve. In the same way, we can calculate the average total cost for each output: the average total cost of producing, for example, the output Os_4 in Diagram 29 is s_4M/Os_4, that is, the slope of the straight line joining O and M. If we calculate the average total cost of each output and transfer our results to Diagram 31, we get the firm's average total cost curve. Both the average variable cost and average total cost curves will be more or less U-shaped, because of the distinctive pattern discernible in the behaviour of the variable and total costs which we described on pages 65–8 above. It will be recalled that this pattern was wholly or mainly a reflection of the Law of Non-Proportional Returns.

If we are to illustrate the firm's sales plan on this new diagram, it is convenient to use a third cost concept, namely, marginal cost. The marginal cost of a unit of output is the increase in variable costs caused by the production of that unit: if, for example, the variable costs of producing 10 units is 20s., and of 11 units is 21s., the marginal cost of producing the eleventh unit is 21s. less 20s., or 1s. Since total costs always exceed variable costs by a fixed sum of money, namely, the fixed costs, the increase in variable costs entailed by producing any additional unit of the product will always be equal to the increase in total costs for the same unit. Geometrically, the marginal cost at any output is measured by the slope of the tangent to either the variable or total cost curves at that output. When we calculate the marginal cost at each output in Diagram 29, and plot our results in Diagram 31, we get the marginal cost curve. The marginal cost curve, like the average variable and average total cost curves, is U-shaped, and it cuts both of these at their minimum points.*

If the firm believes that it will be able to sell its output at a

* Why this must be so can be shown geometrically. The AVC at any output is equal to the slope of the line drawn from the origin to the point on the variable cost curve at that output. In Diagram 29, average variable cost is at a minimum at an output of Os_5, for OB is less steep than any straight line drawn from the origin to the variable cost curve at any other output. At Os_5, the marginal cost is equal to the slope of the tangent to the variable cost curve at B. Since this tangent must coincide with OB, at this output marginal cost and variable cost must be the same. In the same way, it can be shown that at the output Os_3, average total cost is at a minimum and equal to marginal cost.

price of p_1 per unit, then the relationship between expected selling price or average revenue and output is shown by the horizontal straight line drawn at the price p_1 in Diagram 31. Given these expectations, if the firm wishes to maximise its net revenue, it will plan to produce and sell an output of Os_1 per week. This output corresponds to the output Os_1 in Diagram 29. There, at the output Os_1, the slope of the revenue line (which is equal to the expected selling price) is the same as the slope of the total or variable cost curve (which is equal to the marginal cost). In Diagram 31, the equality between marginal cost and expected selling price when net revenue is at a maximum is shown explicitly. We may confirm more directly that at the price p_1 net revenue is greatest when Os_1 is being produced and sold each week. Marginal cost measures the amount by which total cost increases when output is expanded by one unit. The price, given the firm's expectations, is equal to the amount by which the total revenue would increase if sales were increased by one unit. By producing and selling an extra unit of output beyond Os_1, it is clear from Diagram 31 that the firm would add more to its costs than it would to its revenues. By doing so, net revenue would be reduced. Similarly, by producing and selling one unit less than Os_1, more would be subtracted from revenue than would be subtracted from costs. By reducing output and sales by one unit, net revenue would be reduced also. If net revenue would be less at outputs greater and less than Os_1, then it must be at a maximum when Os_1 is being produced and sold in each week.

In the same way, we can show that at an expected selling price of p_2, planned production would be Os_2; at p_3, Os_3, and so on. At expected selling prices of less than p_3 but greater than p_5, the weekly revenue will not cover the weekly total costs. If the contracts made with the fixed factors do not expire and cannot be revoked or revised during the planning period, the firm will continue to operate — planning to sell Os_4 per week at p_4 per unit and Os_5 per week at p_5 per unit — for the reasons stated on pages 70–1 above. The marginal cost curve above the minimum point of the average variable cost curve in Diagram 31 is, therefore, the firm's weekly supply curve of cloth. It is in all respects identical with the supply curve shown in Diagram 30.

It must be emphasised that Diagram 31 adds nothing to

Diagram 29, for all the relationships in it are derived from data portrayed in the latter. The chief danger in using the derived relationships is that frequently they seem to be interpreted behaviouristically. Thus, the sales plan which the firm chooses is commonly described as that which will equate marginal cost and expected selling price. This, however, is merely an alternative way of putting our assumption that the firm strives to make its expected net revenue as great as possible. While to strive to equate marginal cost and price is to strive to maximise net revenue, it is best not to state the firm's objective in this way, for if we do we risk interpreting, or seeming to interpret, the cost and revenue lines on Diagram 31 behaviouristically. Indeed, it is possible that some of the attacks on the 'marginal analysis' (of which our analysis so far is an example) by the proponents of 'Full Cost' or 'Average Cost' pricing spring partly from confusions of this kind, and partly, perhaps, from some confusion about the nature of theory and analysis.

The firm whose weekly supply curve of cloth is illustrated in Diagrams 30 and 31 may not be the only supplier of cloth. We can, however, derive the supply of cloth of each other existing firm that is a potential supplier of cloth of this quality in a precisely similar way. If the prices of the variable inputs are data for all the firms that produce cloth,* then the *total* or *market* supply of cloth may be obtained simply by adding together the supplies of cloth of the firms that are planning to produce and sell it. The way in which this summation is effected is illustrated in Diagram 32. Figures A, B and C show the weekly supply curves of three separate and independent firms. We get the total

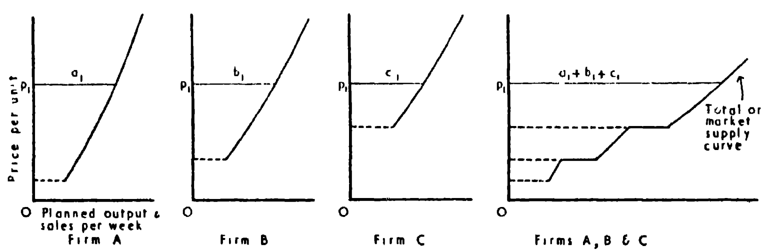


DIAGRAM 32

* If the prices of the variable inputs rise as all the firms that produce cloth expand their outputs (or vice versa), this description of the derivation of the total supply curve must be modified. See *infra* pages 131–3.

supply curve of cloth by adding together the quantities of cloth that each firm would plan to sell at each expected selling price. Thus, at the price p_1, firm A plans to sell a_1 per week, B plans to sell b_1 per week, and C, c_1. The total quantity of cloth that all the firms plan to sell at p_1 is therefore a_1 plus b_1 plus c_1, and this is plotted against the price p_1 in Figure D. In the same way we can discover the quantity of cloth that will be supplied by all these firms at each other expected selling price for cloth. When all these points are joined together in Figure D, we have the total or market supply curve of cloth.

In Diagram 32, it is assumed that the firms A, B and C, have different supply curves of cloth. We would expect this to be generally the case, for there will be differences between firms in the quantity, kind and quality of the productive services that they are using. Different firms may have different production possibilities open to them in the short-run because they made different decisions in the past about the size of plant and the quantity and kind of equipment and machinery to use. There may be differences in the qualities of the variable inputs they use: if each must pay the same time-rate of wages, and if C, for example, because of its location or past behaviour, can employ only the less efficient weavers, then C's costs will be relatively higher and the quantities it plans to sell at each price relatively less, than those of its competitors.

We have assumed that each firm in making it sales plan expects the selling price of its product to be beyond its control. While each firm may plan on the assumption that the prices of its products are beyond its control, the total effect of all firms implementing their sales plans is to assist in the determination of the relations between the prices of the things they sell. The price-determining role of these sales plans is summarised in the total or market supply curve for each product. The manner in which this role is played will be described at some length in Chapter 4.

THE PRICE-ELASTICITY OF SUPPLY

In this chapter so far we have done two things: first, we have shown how the sales plan of an existing firm depends on the production possibilities, on the expected selling price of the product, on the prices the firm expects to have to pay for its variable

inputs, and on the firm's objective; second, we have described the direction in which the contents of the sales plan would alter if there was some change in the expected selling price: the higher the expected selling price the greater will be the planned output and sales, and vice versa. Products may be classified according to the extent to which the quantities of them that are offered for sale vary as a result of changes in their respective selling prices. The responsiveness of planned sales to changes in the expected selling price is called the price elasticity of supply.

The way in which the quantity of cloth, in our example, responds to a change in the selling price is illustrated by the supply curve of cloth in Diagram 30. The strength of the response of quantity supplied to relatively small changes in the expected selling price is measured by dividing the proportionate change in the quantity supplied by the small proportionate change in price that 'causes' the change in quantity supplied. In Diagram 30, if the expected selling price were Op_2, Os_2 would be supplied; if it were to rise to Op_1, the firm would plan to sell Os_1 per week. The price elasticity of supply in this range of price is:

$$E_s = \frac{s_1s_2}{Os_2}\bigg/\frac{p_1p_2}{Op_2}.$$

The numerator, s_1s_2/Os_2, is the proportionate change in the quantity supplied, and we get it by dividing the actual change in the quantity supplied by the quantity supplied before the rise in price. The denominator, p_1p_2/Op_2, is the proportionate change in the expected selling price, and is obtained by dividing the actual change in price, p_1p_2, by the price at which the firm initially expected to be able to sell its product, that is, by Op_2. We can see more clearly how this measure is used by taking a numerical example. Suppose that when the expected selling price is 50 pence per unit, the firm plans to sell 350 yards of cloth per week, and that if the price were to rise to 51 pence per yard, planned weekly sales would be 375 yards. The price elasticity of supply of cloth is then:

$$E_s = \frac{25}{350}\bigg/\frac{1}{50} = 3{\cdot}57.$$

When the value of the price elasticity of supply is zero, we shall say that supply is perfectly inelastic; when it is infinitely large, we

shall describe supply as being perfectly elastic. When the price elasticity is greater than unity, we shall say that supply is relatively elastic; when it is less than unity, supply will be described as relatively inelastic. It can be shown that the price elasticity of supply of a product will not generally be the same at all prices, and that the elasticity in any particular range is not the same as the slope of the supply curve in that range of prices, in a manner analogous to that used on pages 36–7 above. This measure of price elasticity of supply, like our measure of the price elasticity of demand, is only valid if the change in selling price is very small: strictly, it is a measure of point elasticity and should be used only if the price-change is infinitely small.

CHANGES IN SUPPLY

The relationship that we have called supply shows us the sales plan that the firm, in our example, would choose at each expected selling price for cloth, when its production possibilities, its objective, its contractual obligations, and the prices it expects to have to pay for its variable inputs, all remain unchanged. We must now examine what will happen to supply when there is any alteration in one or other of these.

First, the effects of a change in the production possibilities. The production possibilities may be altered by the firm choosing a new method of production, or by extending or contracting its existing buildings and plant while maintaining its existing method. In either case, the isoquant map in Diagram 24 will be replaced by a new one. If the firm's objective and the prices of its inputs remain unchanged, there will be a new weekly supply curve which may bear almost any relationship to the old one. In general, if a firm expands its potential outputs, the new supply curve will usually lie south and east of the old, indicating that the firm will plan to produce and sell more each week at each expected selling price than before.

Second, the effects of a revision in the firm's contractual arrangements with its 'fixed' factors. If these revisions occur at the same time as the firm chooses a new method of production or decides to exploit its existing method differently, then the effects on supply will be those described in the previous paragraph. The only kind of contractual revision that will not alter the range of production possibilities is one which affects only the rewards

of the firm's existing 'fixed' factors. Revisions of this kind will have no effect whatsoever on the firm's weekly supply: provided the quantity and quality of the 'fixed' factors at the firm's disposal remain unchanged, its weekly supply is in no way affected by its fixed costs. A change in fixed costs arising solely from a change in the rewards paid to the fixed factors will, however, alter the length of time for which the firm's existing weekly supply will be maintained. Thus, if the fixed costs were reduced to zero, the firm if it chose could produce indefinitely at prices above minimum average variable cost in Diagram 31.

Third, the effects on supply of a change in the firm's objective. These will depend on what new objective is chosen. The weekly supply, in our example, is what it is because we have assumed, *inter alia*, that the firm wished to earn the maximum net revenue in each period. If the firm wished merely to cover its total costs of production, then its supply curve would be the rising part of its average total cost curve in Diagram 31. If its aim were to earn a constant weekly net revenue, then its supply curve would be a curve lying directly above the rising part of its average cost curve, and asymptotically approaching it as planned output increases.

Fourth, the effect of a change in the price of one or more of the firm's variable inputs. We may ascertain this by repeating step by step the argument of this chapter. If the price of input *Y* falls while the price per unit of input *X* remains the same, then each output can be produced with less expenditure on variable inputs: the expansion path in Diagram 27 will swing towards the vertical axis, and the variable and total cost curves in Diagram 29, the average total cost, average variable cost and marginal cost curves in Diagram 31, and the supply curve in Diagram 30, will all shift southwards and eastwards, for the cost of each output will now be lower. Conversely, if the price of one or other of the variable inputs should rise, the supply curve would shift northwards and westwards: the firm would plan to produce and sell less at each expected selling price than before.

CHAPTER 3

The Sales Plan of the Firm: Long-run

We have assumed throughout the preceding chapter that the firm's current behaviour was circumscribed by past commitments. Some time in the past, the firm built, bought or leased factory buildings of given size and design, installed in them a number of machines and certain quantities of other equipment, and hired managerial and executive labour. While these past decisions still bind it (that is, in the short-run) the firm is limited in each production period to the alternative outputs that these 'fixed' factors can produce with the aid of certain variable inputs. From the range of possible outputs, the firm, in the light of its expectations about the prices of its products and of its variable inputs, chooses that which promises, when produced and sold, to achieve its objective. In the last chapter, we showed also how the going firm would revise its sales plan in response to changes in the expected selling price of its product: the locus of these revisions was the firm's short-run supply curve.

In this chapter, we shall study long-period planning. We shall assume that no past commitments bind the firm: the range of production possibilities and of sales possibilities open to the firm is no longer circumscribed by any fixed factors, for now the quantities and qualities of all inputs can be varied. We shall first, delineate the range of production possibilities open to the firm in this position; next, we shall describe the patterns that have been, or might be, discerned amongst them; and lastly, we shall illustrate the firm's choice of a sales plan, given the expected prices of the product and of the inputs.

LONG-RUN PRODUCTION POSSIBILITIES AND THE PATTERNS IMPLICIT IN THEM

In Chapter 2 (pages 50 ff.), we described the range of production possibilities open to the firm in the short-run. We experienced little difficulty in doing so, for there we assumed that the firm had already decided upon the exact characteristics of its product and on the quantities and qualities of all but two of the inputs required to produce it. In choosing its sales plan, the firm was limited to the different quantities of its product that could be produced by using varying quantities of the two variable inputs, X and Y. It was easy to illustrate the limited range of production possibilities, for there were only two variable inputs, and all of the relatively small units, by which the use of X and Y could be increased or decreased, were assumed to be identical with one another.

In laying plans for the long-period, the problems that face us (and the firm) are more complex. First, the firm is not prevented by past agreement or by lack of time from varying the quantity of any output: the size of the buildings and the quantity of equipment are no longer data. This by itself, however, need create no insuperable difficulty, for the method of analysis of the previous chapter could be extended to problems in which not only two but rather all inputs were variable. Thus, if the firm employed only three inputs, X and Y and Z, all possible combinations of these inputs might be plotted in the three dimensional space bounded by three axes, each at right angles to each other. Beside each point we would write the output which the firm expected that combination of X, Y and Z to produce, and order could be introduced by drawing isoquant surfaces, each passing through all combinations of the three factors that promised the same output. Each isoquant surface would be convex when viewed from the origin, because the inputs X, Y and Z are not perfect substitutes for one another. If we erect a plane at right angles to the axis on which we measure X, for example, we would expect all points on that plane to follow the pattern that was in part explained by the Law of Non-Proportional Returns, for the outputs lying on it are those that can be achieved by combining varying quantities of Y and Z with a fixed quantity of X.

So far our argument is analogous to that of the previous chapter. The analogy breaks down, however, when we seek a pattern in the combinations of X, Y and Z that lie on a straight line drawn through the origin. The Law of Non-Proportional Returns is of no help here, for as we move along such a line the quantities of X, Y and Z are all being increased by the same proportion. Intuitively, we would expect the output to vary directly with the planned input: thus, if $2X+4Y+1Z$ promise an output of 10 units per period, then we would expect 50 per cent more of each input — that is, $3X+6Y+1\frac{1}{2}Z$, to yield an output 50 per cent greater, that is, 15 units per period; and we would expect half as much of each input, that is, $1X+2Y+\frac{1}{2}Z$, to yield only 5 units per period. It is not easy to confirm this assumption in any direct way either by observation or experiment. It may be observed that in many industries there are wide variations in the size of firm; but this does not provide confirmation of our hypothesis that when the quantities of all inputs are increased by a given proportion, output will rise by the same proportion. And this for two reasons: first, if the price of the industry's product is a datum for each firm (as we assumed in Chapter 2), and if all the firms are operating, this means that the marginal cost of the last unit produced in each firm will be the same. However, the average total cost per unit of output might differ widely from one firm to another.* Second, even if firms of different sizes have co-existed in an industry for some time, this at best means that each firm's total revenue at least covers its total costs of production. We might infer that average total cost per unit of output is similar in each firm. But this does not confirm our hypothesis that output varies directly with inputs, for different firms might be using different methods of production and therefore using different inputs. If all firms were using the same inputs, and paying the same unit prices for them, then the fact that average total costs per unit of output was the same for each firm would, of course, support our hypothesis.

Controlled experiments may be equally inconclusive, for it may not be possible to honour the assumptions from which this hypothesis is derived. The experimenter or manager or production planner or co-ordinator might combine different inputs in

* Cf. Diagram 32.

the same proportions (in the above example, $2X : 4Y : 1Z$) but in varying quantities and record the outputs obtained. In such a series of experiments the experimenter's services must be included as a separate input. These services, however, are not being increased by the same proportion as the quantities of all other inputs. Indeed, what is happening is that increasing quantities of all the other inputs are being combined with a fixed quantity — namely, one unit — of management or 'planning'. In these circumstances, we would expect the increases in other inputs to yield more than proportional increases in output for a while, and then less than proportional increases in output: that is, the relationship between output and all inputs other than management will display the pattern described by the Law of Non-Proportional Returns. These results cannot disprove our hypothesis, for the hypothesis describes the relationship that would obtain between output and inputs if *all* inputs were increased by the same proportion, whereas our results were arrived at by varying only all inputs with the exception of management by the same proportion. However, while the hypothesis is not disproved, serious doubts are cast on its usefulness.

If the quantities of all inputs that the firm required, with the exception of management or co-ordination, could be continuously varied, then the analysis of the previous chapter would suffice. The production possibilities open to the firm in the long-run could be illustrated on a diagram which was similar to Diagram 21 and which showed the outputs to be expected from varying quantities of all other inputs when combined with the fixed unit of management: the same patterns could be discerned and the same deductions drawn. Management, however, is not the only input that is incapable of continuous variation. Thus, if a firm uses one motor lorry, it cannot increase the number of lorries at its disposal by less than 100 per cent; if it uses two typewriters, it cannot increase the quantity of this input by less than 50 per cent, or reduce it by less than 50 per cent, for a typewriter must be of a certain minimum size if it is to do its job properly; and if the firm is employing one accountant, it cannot do less than employ another whole accountant. Inputs such as these, the quantity of which cannot be varied continuously with output, are usually called 'indivisible' or 'lumpy' inputs. Top-management or co-ordination is clearly an extreme

example of indivisibility or lumpiness. Another extreme example of indivisibility is the amalgam of fixed factors that the firm has at its disposal during the short-run. The technical consequence of indivisibility is that as more of the other and divisible factors are combined with the indivisible factors, output follows the pattern described by the Law of Non-Proportional Returns.

Most factors that a firm uses are indivisible to some extent. The quantity of the factor may be incapable of continuous variation for technical reasons, as with typewriters and lorries, for each of these must be of a certain minimum size if it is to do the work for which it was designed. The indivisibility might arise for reasons that are partly technical and partly institutional: thus, the firm might not be able to hire labour-service in units of less than one hour or one week. Whether or not the degree of indivisibility merits the adjective 'indivisible' depends mainly on the number of units of that factor that the firm is using. If a firm is engaged in road haulage and if it is already operating 200 lorries, then the degree of indivisibility in the input lorries is unlikely to be important. If the firm is a small wholesaler owning only one lorry, then the degree of indivisibility will be significant.

The notion of indivisibility depends also on the units in which we measure inputs. The input transport services may be measured in number of lorries or in ton-miles: indivisibility is more likely to be significant if we use the former than if we use the latter. The input typing services may be measured in numbers of typewriters or in words typed per hour: the degree of indivisibility may be the less worthy of note if we use the latter units. In general, with durable goods (like lorries, machines or buildings which yield their services over many production periods), indivisibility will appear more important if we measure inputs in terms of the number of such goods rather than in terms of the services which they render. This choice of units is rather more than a linguistic quibble: a firm cannot have one-half of a lorry for one week, but if a lorry gives 100,000 ton-miles per week, a firm can have 50,000 ton-miles of input by hiring a lorry for three days, or it may procure the same quantity of input by having another firm transport its goods. If a firm has more work than one accountant can cope with but less than two could do, then it may hire accounting services from a specialist firm.

We conclude, therefore, that indivisibility of the kind we have been describing is seldom important: it is always a matter of degree and frequently the result of how we define input. Further, a firm can generally vary the quantities of the services of durable goods that it uses, more or less continuously, by hiring these from other firms. It is possible that the only example of inescapable indivisibility in the long-run is the factor that we have called top-management, co-ordination, planning or entrepreneurship. By its very nature one unit, and one unit only, of this is required by each firm.

We have thus far been assuming implicitly that the firm would plan to produce higher outputs with precisely the same factors (though with a larger number of units of each of them) as it would use to produce smaller outputs. This, however, is neither necessarily nor generally the case. We observe that in industries, where firms of different sizes co-exist, the small firms use different inputs from the large firms, because they use different *methods* of production. The labour-service and machinery employed by small firms are generally less specialised than those employed by large firms. When output is large it is possible to keep one man fully occupied on a narrow range of tasks, or even on a single task such as tightening one particular screw; when outputs are small, one man may have to do many jobs. Thus, when the firm is small, the owner or manager may be responsible for buying raw materials and hiring labour, selling the product, organising production, keeping an eye on costs, and co-ordinating all these decisions to achieve the firm's objective; when the firm is large, the functions of buying, selling, producing and costing may each be large enough to keep a specialist fully occupied and the manager remains ultimately responsible for policy. We observe also that as the size of firms (as measured by output) has grown, specialised machines have been developed to do jobs that were formerly done by workers.

The factors that are used may vary with the size of the firm, not only because they become more specialised, but because they increase in size. Thus, a small firm may obtain its power from one or two small electric motors; a large firm may use electric motors of much greater horse-power. A small firm may have small buildings and a small boiler in its engine-room; the large firm, large buildings and a large boiler. And these large durable goods

perform the same functions as the smaller ones of the same name. We shall see later that the larger firms in any industry use bigger and more specialised inputs — that is, use different methods of production — than the smaller firms, because such methods are less costly.

We may sum up thus far as follows: In making a long-run decision, the firm is choosing the framework with which it will have to operate in the ensuing short periods — for once the firm has decided upon the size of building and the number, kind and quality of machines to use, and so on, it will be confined for some time thereafter to the range of outputs that these can produce when used with the appropriate 'variable' factors. At the time when the firm is making its choice, it may draw on the existing fund of technical knowledge to discover the different methods by which any particular range of outputs, of the genus of product in which it is interested, might be produced. Thus, it may be that the range of output, 1,000 to 2,000 units per week, could be produced in a building of size and type *A*, with 20 machines of type *B*, when fed by varying quantities of a specific quality of raw material *C*, and operated by differing quantities of a certain grade of labour-service *D*; or the same outputs might be achieved in the same building with fewer machines of a different type, and with different quantities and qualities of labour-service and materials; or, indeed, with any one of an almost infinite number of different combinations of different 'fixed' factors, each set to work by varying quantities of the appropriate 'variable' factors.* For each range of output, then, we may draw a number of diagrams, each similar to Diagram 21, and each assuming some given bundle of 'fixed' factors and showing the different outputs that can be obtained from that bundle with varying quantities of the 'variable' inputs appropriate to it.

The total of all such diagrams for all ranges of output illustrate the production possibilities open to the firm in the long-run, and the previous pages indicate their main characteristics. With each method of production — that is, with the use of the

* Before the firm makes its choice, all (or almost all) factors are potentially variable. After the firm has made its choice, however, some are fixed and some remain variable. In delineating the production possibilities, we use 'fixed' to mean those factors that would be fixed, and 'variable' to mean those factors that would remain variable, were the firm to make that particular choice.

same inputs — the output can be increased by using more of the same kind and quality of the 'fixed' and 'variable' factors. Some of the 'fixed' factors may be more or less indivisible, but we shall assume that the degree of indivisibility is only significant for top-management. The long-period production possibilities that are open to the firm within any given method of production are illustrated roughly in Diagram 33. On the vertical axis, we measure output per period and on the horizontal axis we measure the quantities of the 'fixed' and 'variable'* inputs that are being combined with the given unit of co-ordination. Curve 1 shows us the outputs per period that the entrepreneur might expect if he chose bundle 1 of the 'fixed' factors and used them with increasing quantities of the 'variable' factors. And similarly, curves 2, 3, 4, and so on. Each of these curves will be of the same shape as that shown in Diagram 25, and for the same reasons: each shows the behaviour of output as some variable factors are combined with a bundle of 'fixed' factors and the given unit of entrepreneurship. Not only does each of these curves follow a particular pattern: taken together, they follow the

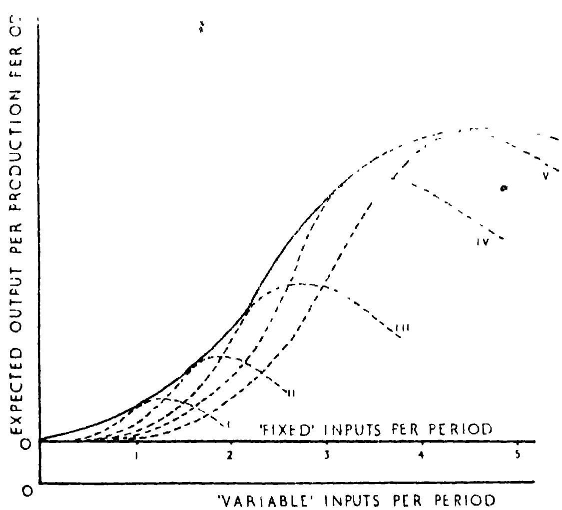


DIAGRAM 33

* If there is only one 'variable' input in this method of production, no problem arises in measuring it on the horizontal axis. If there is more than one 'variable' input, unambiguous measurement is only possible if they are all used in fixed proportions with one another.

same general pattern (which is illustrated by the thick, unbroken line in the diagram), and for the same reasons, for as output is increased, more of both the 'fixed' and 'variable' factors are being combined with the given unit of planning or entrepreneurship.

The production possibilities open to the firm in the long-run, were it to use any other method of production, can be illustrated in the same way. For each method of production, we will have a thick line similar to that in Diagram 33, summarising the relationship between output per period and the quantity of 'fixed' and 'variable' factors required by that method of production. All these diagrams will summarise the production possibilities open to the firm in the long-run. The problems facing the firm are first, to decide how to produce each range of output and then to decide what range of output to produce and sell. Neither of these choices can be made on technical or physical grounds alone, for the diagrams illustrating the production possibilities open to the firm within each method of production cannot be compared with one another. And this for two reasons. First, while the genus of product is the same for all methods of production, its species will vary from one method to another. On the vertical axes we may be measuring cotton cloth or motor-cars, but the precise quality of the cloth and the precise kind of motor-car will not be independent of the method of production by which it was made. To take an extreme example: cloth or cars made by the handicraft method of production will not be the same as those that are mass-produced. Second, the inputs that are being measured on the horizontal axis are not the same for each method of production. The 'variable' inputs may vary from highly skilled labour that plays a major role in shaping the product to unskilled or semi-skilled labour that merely tends the machines that make it, and the size and function of the machines will vary also.

These problems are similar to those which we met in Chapter 2. There, the production possibilities open to the going firm were limited by the ridge lines on Diagram 21. The firm was faced with two decisions: first, what quantities of the variable inputs it would use to produce each weekly output — that is, 'how' to produce each output, and second, what particular output to produce and sell. Neither choice was possible on technical grounds, for the different combinations of the two variable

inputs, X and Y, by which each output might be produced could not be directly compared with one another, since there was no common, physical unit to which both X and Y could be reduced. In the long-run, as in the short-run, these choices can only be made in the light of the economic data, namely, the price per unit of each factor of production that the firm might use, the price at which it expects to be able to sell its product, and the objective towards which it is striving.

THE CHOICE OF HOW TO PRODUCE EACH RANGE OF OUTPUT

The production possibilities show us the different combinations of the same or different factors by which each range of output may be produced. The firm must choose what range of outputs to produce and the precise method by which to produce it. The latter choice can be made when the prices of all inputs are known, for when the prices of all the factors of production are given the money cost per unit of output can be calculated. We shall assume that the firm will always plan to produce each range of output by using that combination of inputs that promises the lowest cost per unit.

Let us begin with curve 2 in Diagram 33, showing the alternative outputs that can be produced in each production period with bundle 2 of the 'fixed' factors when worked by the 'variable' factors, and let us suppose that bundle 2 consists of one factory building of given size and design, ten identical machines and one works manager, and that the variable factors are operative labour and raw materials of given qualities and kinds. The price per unit of each of these is given and the entrepreneur's problem is to calculate the behaviour of cost per unit of product over this range of output. Any output within this range may be produced and sold within each production period — which we called one 'week' in Chapter 2. The 'variable' inputs are bought and used within each production period, so that their cost can be attributed wholly to the outputs they help to produce in that period. With the 'fixed' factors, however, the problem is more complex. The 'fixed' factor may be a durable good like a factory or a machine. A durable good is a reservoir of productive services that may be tapped as required: a machine, for example, will yield productive services in each of the production periods that

together make up its life-span. The entrepreneur's problem is to place a value on the machine's contribution in each period, and in solving it three things are relevant: namely, first, the price of the machine; second, the length of its life; and third, the way in which the price of the machine is to be allocated to each of the production periods making up its life.

We shall suppose that the first of these is a datum to the firm. The second, the life of the machine, is a matter for conjecture. If the machine is used continuously, there may come a time when the costs of repairs and replacements in each production period would exceed that period's share (however calculated) of a new machine. The entrepreneur may base his estimate of the life of the machine on when he expects this to happen. The economic life of the machine, however, may be much shorter than the time taken for it to 'wear out'. If an industry has become accustomed to continuous changes in the techniques of production, entrepreneurs may expect this to continue in the future, and each, when contemplating the purchase of a machine of the current type, may base his calculations on the time he expects to elapse before it becomes obsolete. It would clearly be quite fortuitous if the time that is expected to elapse before the machine becomes obsolescent is the same as that required for it to wear out. When making his calculation, the entrepreneur will choose whichever of these periods is the shorter. In older industries, like cotton whose technological experiences have for a while been placid, we would expect the life-span of a machine to depend mainly on the time needed to wear it out; in newer industries, like plastics, chemicals, radio and television, obsolescence almost certainly exerts the stronger influence.

Knowing the price of the machine, and having decided upon its expected life, the next problem is to recoup its price from the production periods over which the entrepreneur expects to use it. The art of accounting provides a number of conventional solutions to this problem, but it is no part of our purpose either to describe or assess them. The entrepreneur will choose one of these, and having done so, he can calculate the share of the cost of acquiring the machine that is to be attributed to each of the production periods that together make up its life-span. This sum of money will be a constituent of the firm's 'fixed' costs for that period. For the sake of simplicity, and not because it is

necessarily the most prevalent convention, we shall suppose that the costs of acquiring the machine or building are distributed equally over each of the periods in which he plans to use it. Thus, suppose that an entrepreneur borrows £100 to buy a machine with an expected life of 10 production periods, and contracts to pay 10s. (that is, ½ per cent) per period as interest: on our assumption the 'fixed' costs of the machine in each period will be £10 10s. 0d.*

We have assumed so far that the entrepreneur is planning to buy the durable good outright. He may, however, plan to rent, lease or hire its services from another firm. If a factory building is leased at a fixed annual rent, or a works manager hired on a six-monthly contract at a stated salary, we shall suppose that each production period is expected to bear an equal share of the rent or salary. If the entrepreneur who leases the factory is responsible for its maintenance, we shall suppose that these costs are also equally distributed over the production periods. Once the cost of acquiring outright or leasing, renting or hiring each of the factors that make up bundle 2 in Diagram 33 is given, the part of these charges that is to be attributed to each production period can then be calculated; on our assumption these 'fixed' costs will be the same for each period. There remains one fixed factor — the entrepreneur — whose remuneration is a part of the fixed costs. We shall value his services at the minimum sum which he must receive in each period to induce him to produce this product; this 'minimum sum' will be equal to the maximum income that he believes he could earn for his services in any alternative use.†

When the expected fixed costs are known for each period, we can proceed as we did in Chapter 2 and calculate the variable costs of each of the outputs lying on curve 2, and by addition obtain the total costs of production of each output. From these, we can obtain the average total cost per unit of output, for in illustrating the choice of how to produce each range of output it will be convenient to work in terms of averages rather than in

* Strictly, it will be rather less than this, for the sum set aside towards repayment of the principal in period 1 can be lent to another entrepreneur during periods 2 to 10 inclusive, and it will earn ½ per cent per period. And similarly for the sums set aside in the ensuing period (see *infra*, Chapter 5).

† For a fuller treatment of the valuation of the entrepreneur's services, see *infra*, Chapter 10.

terms of totals. From the relationship between output and inputs portrayed by curve 2 in Diagram 33, we can get, then, a relationship between output and cost per unit of output: that is, an average total cost curve like that in Diagram 31.

By proceeding in precisely the same manner, and knowing the prices of the relevant inputs, we can get an average total cost curve for each other curve in Diagram 33. All these average cost curves are graphed in Diagram 34. As the planned output is increased within this method of production the average total cost per unit of output falls for a while and then begins to rise: thus, in the diagram, the minimum points of the average cost curves are progressively nearer to the horizontal axis until bundle 3 of the 'fixed' factors is being used; when greater quantities of the same 'fixed' factors are used, average total cost begins to rise. The average cost curves display this pattern because of the pattern followed by the dotted line in Diagram 33, and that in turn was explained by the Law of Non-Proportional Returns, with entrepreneurship as the fixed factor with which increasing quantities of the 'fixed' and 'variable' factors were being combined.

If this were the only method by which the product in question could be produced, and if the entrepreneur were contemplating the production of outputs between *Oa* and *Ob* per period, he would decide to do so with the bundle 1 of the 'fixed' factors, for

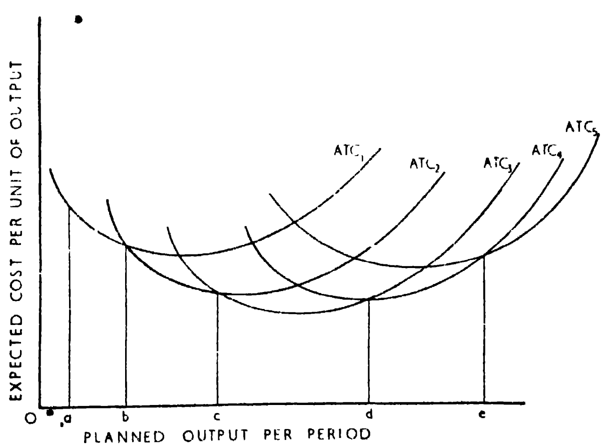


DIAGRAM 34

at the given prices for inputs this promises the lowest cost per unit. Any other bundle of the 'fixed' factors will give higher costs per unit in this range of output: between *a* and *b* all points on ATC_1 are below ATC_2 or ATC_3 or any other average total cost curve.* Similarly, the entrepreneur would plan to produce outputs between *Ob* and *Oc* per period with the bundle of 'fixed' factors that promises the unit costs shown by ATC_2, and outputs between *Oc* and *Od* with the factors that would give ATC_3, and so on.

The entrepreneur, however, is not usually limited to one method of production: there are generally a number of them from which he may choose. For each other method of production, we can draw a diagram similar to Diagram 33, and once the prices of the inputs that are required with that method are given, average total cost curves can be drawn as in Diagram 34; they will follow the same pattern and for the same reasons. All these diagrams can be superimposed on one another, for the vertical axis of each measures sums of money and the horizontal axes measure units of the product per period.† When this is done, we have a picture of the cost possibilities open to the firm when planning for the long-run. The entrepreneur will plan to produce each range of output by the method of production and with the quantities of the inputs demanded by that method that promises the lowest average cost per unit of the product. These choices are illustrated in Diagram 35, where we suppose that there are only two methods of production — the *ATC* curves drawn in unbroken lines relate to the first and the curves drawn in dotted lines relate to the second method. For outputs per period of less than *Ox* some quantities of the inputs appropriate to the first method will be chosen; for outputs greater than *Ox* the second method will be chosen.

In this diagram, we have implicitly assumed that one method of production will give lower costs for some or all outputs than another. We must now attempt to explain why this may happen. It will be remembered that we have defined a method of pro-

* As drawn, the diagram assumes that the 'fixed' factors appropriate to this method of production are not infinitely divisible. If they are, then there will be an infinite number of these *ATC* curves. They will all, however, follow the pattern we have described.

† This is not strictly true. On the horizontal axis we are measuring the same genus of product, but its species probably varies from one method of production to another. For simplicity's sake, we shall ignore this complication at this stage.

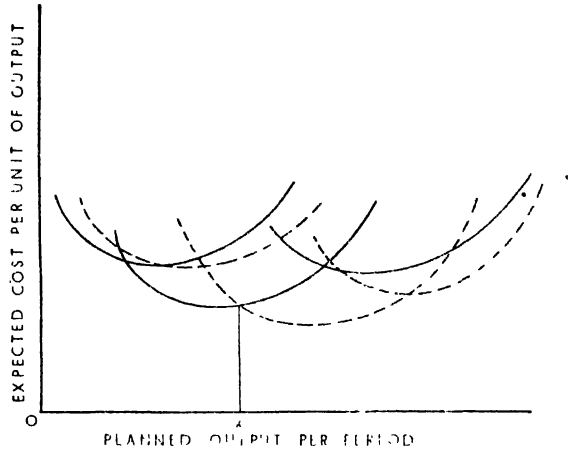


DIAGRAM 35

duction in terms of the factors that are used: different methods require different factors. The factors may differ in size, while performing the same productive service: thus, power may be supplied by electric motors of widely different horse-power, or the input floor-space may be supplied by factory buildings of different sizes. The factors may differ not merely in size but also in kind, while performing the same or similar functions: thus, a firm may obtain transport services by hiring a boy with a bicycle, by using a horse and cart or by buying a motor-van; an executive may hire a skilled shorthand-typist or buy a dictating machine and employ a typist; earth may be moved by many unskilled workers with picks, shovels and barrows or by highly specialised earth-moving machinery and a few skilled operatives.

We should not be surprised that different methods promise different unit costs of production over any range of output, for different factors will have different prices. The fact that methods using large or specialised durable goods promise lower costs over some ranges of output than those requiring small or more versatile equipment, must then be due to the relationship between the prices of the two kinds of equipment. We normally find that the price of machines does not vary directly with their size and capacity, and we may illustrate how this affects costs of production by a simple example. Let us suppose that in some industry a power input of one horse-power is required to pro-

duce 10 units of the product in each production period, and that power can be obtained from two sizes of electric motor — 1 h.p. and 1,000 h.p. From the technical point of view, if an output of 10,000 units in each period were contemplated, the power might be obtained from either one thousand 1 h.p. motors or from one 1,000 h.p. motor. Ignoring operating costs, if the price of a 1,000 h.p. motor were 1,000 times that of the 1 h.p. motor, and if the expected life of each size of motor were the same, then the fixed costs per period of this input would be the same. In practice, however, the price of the larger motor will probably be only 100 times greater than that of the smaller, so that its fixed costs per period will be only one-tenth that of the 1,000 smaller motors; other things being equal, therefore, the average total cost will be lower when 10,000 units are being produced with the aid of the single large motor than with a large number of small ones. Thus, if operating costs were zero and electric motors had a life of only one period, and if the price of a 1 h.p. motor were 100s. and that of a 1,000 h.p. one 10,000s., then the 'fixed' cost per unit when 10,000 units of product were being produced would be 1s. when the large motor was used and 10s. when 1,000 of the small ones were employed. At smaller outputs, however, the larger motor might be more expensive. Thus, if an output of 100 units per period were being contemplated, the 'fixed' cost per unit of product would be 100s. with the 1,000 h.p. motor (assuming that it was technically possible to use it intermittently) and 10s. per unit with ten 1 h.p. motors.

It is sometimes said that large durable goods are more efficient than small ones that do the same job and their greater 'efficiency' is attributed to 'indivisibility'. This is allowable provided that by 'efficiency' we mean (which we do not always mean) cost per unit of product, and by 'indivisibility' the fact that a 1 h.p. electric motor, for example, costs more than the one-thousandth part of the price of a 1,000 h.p. motor. It is perhaps better, however, to eschew these terms in this context, for efficiency is best reserved for its technical use to mean the relationship between input and output, and indivisibility to the meaning we have given it on pages 84–6 above. We might speak of the phenomenon described in the previous paragraph as 'cost indivisibility', but there seems little to be gained by doing so.

Many other examples may be listed to illustrate this fact that

as the size of a durable good is increased its price, for a while at least, rises less than proportionately. Thus, a 2,000 ton tanker will not cost twice as much as a 1,000 ton tanker; the price of a boiler with a capacity of 1,000 cubic feet will be less than twice that of a boiler with half that capacity, and one two-pound glass jam-jar will be cheaper than two one-pound jam-jars. We shall describe in Chapter 7 how the relative prices of the factors of production are determined. At this stage, it suffices to note that costs of production help to determine price and that these costs do not vary directly with the size of the good that is being produced. Thus, to take the simplest example, a wooden box with the dimensions 3 ft. × 3 ft. × 3 ft. will have 27 times the capacity of one measuring 1 ft. × 1 ft. × 1 ft., but only 9 times as much wood will be needed to make it: if the two boxes can be made of the same wood of the same thickness, the material costs for the larger box will tend to be only nine times that of the smaller. And the same argument may be used to explain why large ships or large boilers are proportionately less expensive than small ships and boilers. Beyond a certain size, however, price may rise more than proportionately. Thus a box of 5 ft. × 5 ft. × 5 ft. may be made with $\frac{1}{4}$-inch timber nailed together, but a box of 10 ft. × 10 ft. × 10 ft. may have to be made from $\frac{3}{4}$-inch timber morticed together, if it is to do the same kind of job. The larger box has a capacity 8 times greater than the smaller, but the volume of wood required to make it is 12 times greater ($37\frac{1}{2}$ cubic feet as compared with $3\frac{1}{8}$ cubic feet), and if timber is sold at a given price per cubic foot, its cost will be 12 times greater also. In this example, the larger box will give a higher 'fixed' cost per cubic foot of packaging capacity per period than would 8 small boxes. The same will be true of ships, boilers, buildings, and so on.

In the last few paragraphs, we have confined our attention to the effect of size on unit costs. The same arguments can be used to show why more specialised machines or labour promise lower costs at larger outputs. A trench may be dug by unskilled labourers with shovels or by a mechanical ditch-digger operated by a few skilled workers: the relationship between the prices of these factors in most Western countries is such that if the trench is a long one and the terrain suitable it will be less costly per cubic foot to excavate by the latter method. The explanation of

why one method of production gives lower unit costs at some outputs than others must lie in the relationship that obtains between the prices of the inputs appropriate to the different methods, and we shall describe what determines these relationships later, in Chapter 7.

We have now described the long-run production possibilities open to the firm and the patterns that may be discerned amongst them. Given these, and given also the price of each factor and the entrepreneur's expectations about the life-span of durable goods, the average cost per unit of output can be ascertained, for all methods for all outputs. Within each method of production, average costs will fall for a while as planned output rises and then begin to rise; this is mainly due to the Law of Non-Proportional Returns operating with management as the fixed factor. If we arrange the methods of production from which the entrepreneur may choose according to their age — that is, the oldest first and the newest last — we will generally find that the newer methods offer lower costs per unit at larger outputs. And this should not surprise us, for the incentive to improve the techniques of production was the urge to lower costs. If the average total costs of production are calculated for each range of output when produced by each method of production, the range of choice will appear as in Diagram 36. Given any particular range of output, we shall assume that the entrepreneur

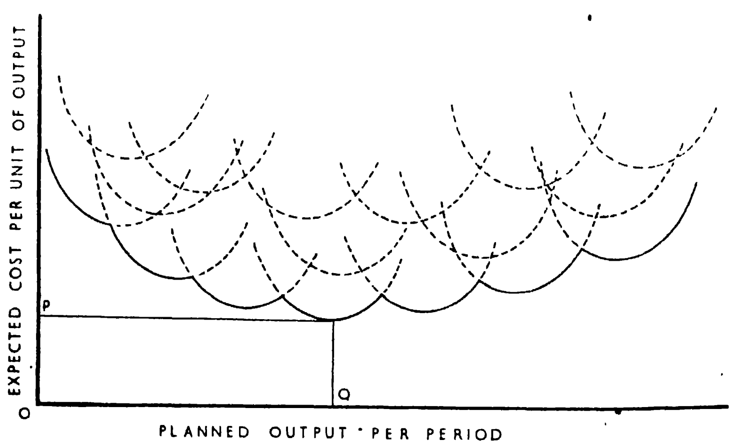


DIAGRAM 36

will plan to produce it by that method of production and the quantities of the factors required by it, which promises the lowest average total cost per period. He will therefore restrict his choice to the cost possibilities lying on the thick black line in the diagram. This line is called the Long-Run Average Cost Curve or Planning Curve. Having decided *how* to produce each range of output, the next problem is to decide which particular range of output to produce and sell, and this choice is made in the light of the entrepreneur's expectations of the selling price of the product and the objective he is pursuing.

THE CHOICE OF A SALES PLAN

In Diagram 36, each of the average cost curves of which a part helps to make up the Planning Curve is the short-run cost curve that the entrepreneur would have if he actually chose the bundle of 'fixed' and 'variable' factors to which it relates, and if his expectations about the output to be obtained from these inputs and their prices were fulfilled. If we are given the price at which he expects to be able to sell his product in each of the ensuing production periods, and if we assume that he desires to earn the maximum net revenue per period, we can determine which he will choose quite simply. For each of the average total cost curves that makes up the Planning Curve he can calculate what output he would plan to produce and sell were he to choose it, and so discover what net revenue to expect in each period. Thus, if he should choose the average cost curve drawn in Diagram 31, and if the expected selling price were Op_1 per unit, he would plan to produce and sell Os_1 per period and hope to earn a net revenue of $RLMp_1$ per period. Having made this calculation for each 'short-run' cost curve, he will decide on that which promises the greatest net revenue per period. The plant, etc., which promises this will also, on our assumptions, promise the greatest aggregate net revenue over all the production periods that must elapse before a new long-run decision is made, for we have assumed that the expected selling price of the product, the 'fixed' costs and the prices of the 'variable' factors are the same in each period. At expected selling prices greater than Op in Diagram 36, he will decide to have an average total cost curve that lies to the right of Q. At selling prices of less than Op, he will decide not to produce this product.

The choice of a long-run sales plan can be illustrated diagrammatically. In Diagram 37, we assume that there are no marked indivisibilities either of factors or methods, so that the planning curve is a continuous line. In Diagram 36, there were a finite number of alternative short-run average cost curves each contributing a segment to the planning curve; if there were an infinite number of such curves, the segment that each contributed would contract to a point — the point that showed the minimum average total costs of production for a particular output. If all such points were joined together, we would have a planning curve like that in Diagram 37. On this diagram, we have also drawn the long-run marginal cost curve: it will bear the same relationship to the long-run average cost curve as does the short-run marginal cost curve to the short-run average cost curve. If the firm expects the price to remain at OP_1 indefinitely, it will plan to produce and sell an output of OM_1 in each period. The proof of why its expected net revenue per period will be at a maximum at this price and output is identical with that used on pages 74–5 above, where we showed that expected net revenue per period in the short-run will be greatest when price and short-run marginal cost are equal. When this decision is implemented, the short-run marginal cost of producing the OM_1-th unit per period in that plant will be equal to the long-run

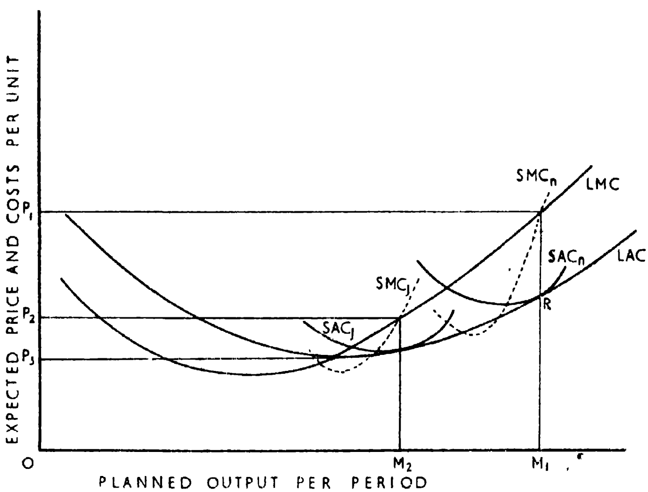


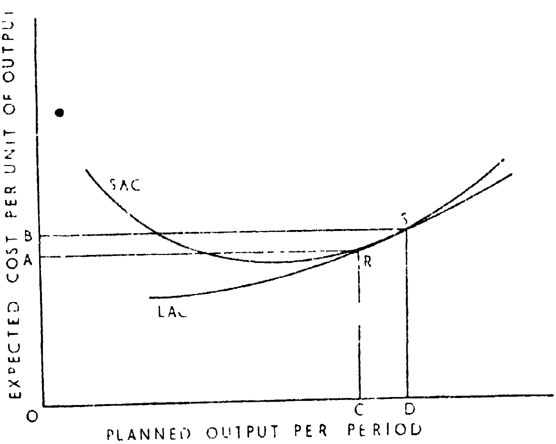
DIAGRAM 37

marginal cost of production of that unit: that is, in terms of the diagram, SMC_n will intersect LMC vertically above R, the point where SAC_n touches LAC.*

In the same way, we can show that if the firm were to expect the price OP_2 per unit to rule indefinitely, its expected net revenue would be greatest if it produced and sold OM_2 per period, with the bundle of 'fixed' and 'variable' factors that would give the short-run curve SAC_7. If the price that the entrepreneur expected to get was less than OP_3 per unit, he would not produce this product, for at such prices the total revenue would fall short of the total costs so that net revenue would be negative. It should now be clear that the firm's long-run supply curve will be that part of its long-run marginal cost curve that lies above the minimum point of its long-run average cost curve. The long-run supply curve of the product can be obtained by adding together the long-run supply curves of all the firms that might plan to make it, if the prices of all inputs are data for all the firms.† Its derivation could be illustrated on a diagram similar to Diagram 32.

PLANNING FOR THE INTERMEDIATE PERIOD

In this chapter so far we have assumed that the entrepreneur in making his plans was bound by no past commitments what-



* To make the proof less mathematical, suppose that SAC_n and LAC coincide at R, not at a point, but over a range of one unit of output, shown by RS. The short-run and long-run total costs of producing OC per period are shown by the area $OCRA$, and of producing OD per period by the area $ODSB$. The long-run and short-run marginal costs are then each equal to $ODSB$ less $OCRA$.

† See footnote on page 76 above.

soever. In the previous chapter, his range of choice was so drastically restricted by past commitments that only a few inputs remained variable. We have chosen these two extremes to focus attention on the importance of time: in general, the longer the period for which the entrepreneur is planning the wider the range of choice that is open to him, and vice versa. In practice, however, many firms plan for periods that lie between our long-period and our short-period, and we shall deal briefly with these intermediate period plans. For a firm, as for all of us, time cannot be divided into discrete periods: rather, it is a *continuum*. All decisions create an 'envelope' within which future choices are confined; the envelope is larger and less confining the longer the period to which the decision that creates it relates and the longer the period for which it is binding. Thus, in our terminology, whether the egg will be poached or boiled for breakfast is a short-run decision; the choice of curtains or carpets is an intermediate period decision, and marriage or buying a house is a long-period decision.

It is only infrequently that an entrepreneur will make a long-run plan of the kind that we have described. Having implemented it, however, some revisions may still be possible. A firm, for example, may have chosen the group of factors that promise SAC_n in Diagram 37. Soon after this choice has been made, there may be a new invention, which shifts the minimum range of the planning curve south-westwards. If this choice had been known when the firm was making its original choice, it would have chosen differently. What the firm will do will depend on the relation between the costs per period it now has, and those it could have had were it now free to choose. If the expected net revenue per period with the latter exceeds the 'fixed' costs plus the net revenue in each period with the existing method, then it may scrap the existing plant, etc., and start anew. New developments, however, seldom have such drastic effects. They are frequently such that they can be used in existing plants and offer some reduction in costs per period — though smaller reductions than would have been achieved if the plant had been initially designed to make use of them. When a firm is deciding whether or not to install improved machines, for example, its choice can be illustrated in a manner similar to that described above. We could delineate the range of production possibilities

open to the firm when it is planning for the intermediate period: this will be narrower than in the long-run for some factors are now fixed, but wider than in the short-run for more inputs are now variable. When the expected prices of the factors are given, these production possibilities can be translated into alternative short-run cost curves that the firm might choose, and it will decide on that which promises the greatest net revenue per period. If the firm decides to instal the new machines, then its new average total cost curve will lie neither on the old planning curve nor on the new one but somewhere between the two.

CONCLUSION

In Chapter 2, we assumed that the firm's behaviour in each production period was circumscribed by certain fixed factors. In this chapter, we have attempted to show why the firm at some time in the past chose those factors in these quantities. We began by describing the physical production possibilities open to the firm that is bound by no past commitments, and showed that each range of output could be produced by many alternative combinations of the same or different inputs. When the price that the firm expects to have to pay for each input is known, the range of choice facing it can be narrowed, for we assumed that the firm would plan to produce each range of output by the combination of inputs that cost least. The final choice of what range of output to produce and sell depends on the firm's objective and on the expected selling price of its product.

In the next chapter, we shall describe how the relative prices of the things that firms sell are determined by the interaction of the sales plans of firms and the purchase plans of households. In doing that, we shall be combining the analyses of this and the last two chapters.

CHAPTER 4

The Determination of Relative Product Prices

In Chapter 1, we described the derivation of a household's demand for any good that it might plan to buy. A household's demand for some good (say, a particular kind of cloth) is a schedule which shows us how its purchase plan would be revised if the only planning datum that altered were the price it expected to have to pay per yard of cloth: that is, the demand for cloth is a schedule that shows the quantity of cloth that the household would plan to buy in a given period of time at each price at which cloth might be sold, *ceteris paribus*. The *cetera* that must remain *paria* are the household's tastes and preferences (that is, its indifference map), its planned consumption expenditure, the prices of all other goods that it might buy, and its desire to obtain the maximum satisfaction from its expenditure. The total or market demand for cloth is obtained by adding together the demands of all the households in the economy that might plan to buy it.

In Chapter 2, we derived the supply of cloth of an existing firm. The firm's supply of cloth is a schedule that shows us how its sales plan would be revised during the short-run if the only planning datum that altered was the price at which it expects to be able to sell its cloth: that is, it gives us the quantity of cloth that the firm would plan to offer for sale in each production period at each price, *ceteris paribus*. The *cetera* that must remain *paria* are the firm's production possibilities (that is, its isoquant map), the prices at which it expects to be able to buy its variable inputs and the objective that it is pursuing. The total or market supply schedule of this kind of cloth is obtained by adding together the supplies of all the firms in the economy that might plan to sell it.

The total demand for cloth summarises the role that households play in determining the relative price of cloth as they implement their plans to buy it. The price-determining role of

firms is summarised in the total supply curve of cloth. In this chapter, we shall: first, describe how these roles are played; second, examine some of the applications of demand and supply analysis in order to demonstrate its usefulness, and third, make use of the analysis in Chapter 3 to analyse price determination in the long-run.

PRICE DETERMINATION: SHORT-RUN

In Diagram 38, we measure the expected price per yard of cloth on the vertical axis, and on the horizontal axis we measure the planned sales of cloth by firms and the planned purchases of cloth by households in each period of time. The market demand and supply schedules are graphed between these axes. The price of cloth will tend towards the level *OP*, for only at that price will the quantity of cloth that firms plan to sell (*OQ*) be the same as the quantity that households plan to buy (*OQ*) in each period.

We can see clearly from the diagram that *OP* is the only price at which the plans of households and firms will be consistent with one another. Thus, if the price were *OA*, firms would plan to sell *OH* during the period but households would plan to buy *OJ*. If firms actually offer for sale an amount equal to *OH*, then the purchase plans of the households must be under-fulfilled by *HJ* during the period. Conversely, if the price were *OB*, households would plan to buy only *OK*, while firms would plan to pro-

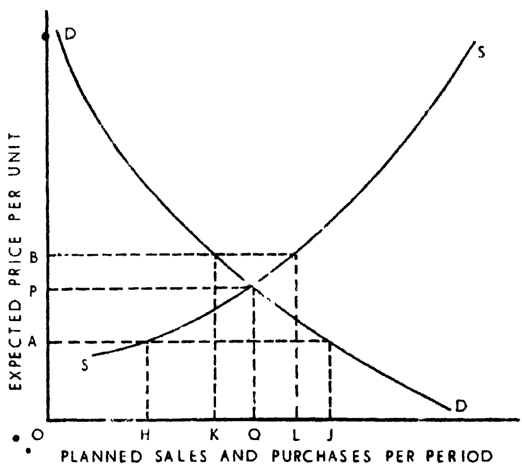


DIAGRAM 38

duce and sell OL. If both the households and the firms attempt to make their plans effective during this period, then the sales plans of the firms will be under-fulfilled by KL — that is, at the end of the period, they will be left with unsold stocks of cloth equal to KL yards. These divergences between the planned and actual purchases of households or between the planned and actual sales of firms cannot continue, and we shall describe presently how their existence sets in motion forces that will probably lead to this commodity being bought and sold at OP per unit.

The price OP per unit is called the *equilibrium* price, and the price will remain at that level, with an even flow of sales and purchases each equal to OQ in each period, so long as there is no change in the demand for cloth or in the supply of it. We showed in Chapter 1 (pages 31–3) that the relationship between price and planned purchases that we called demand will alter if there is any change in households' tastes and preferences, their planned consumption expenditures, their objectives, or in the price of any good (other than this kind of cloth) that they might buy. If the preferences for this kind of cloth become stronger, or planned consumption expenditure increases, or the prices of other kinds of cloth rise, then households will plan to buy more at each price than before. This increase in demand is shown in Diagram 39 by a movement of the demand curve from D_1D_1 to D_2D_2. If there is no change in supply then the price of this grade of cloth will

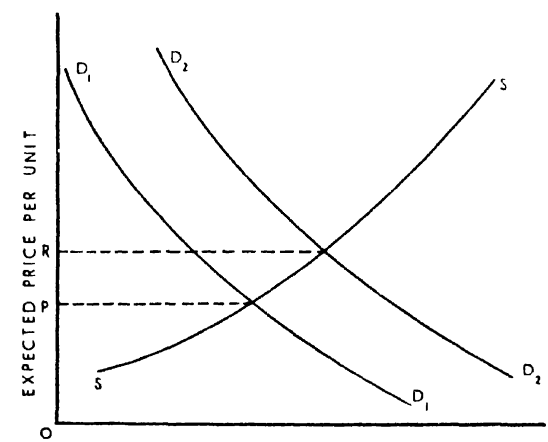


DIAGRAM 39

tend to rise from *OP* to *OR* per yard. The rise in price that follows any given increase in demand will be the greater the less is the price elasticity of supply, and it will be the less the greater is the price elasticity of supply.

We showed in Chapter 2 (*supra*, pages 79–80) that the relationship between expected selling prices and planned sales that we called supply will alter if there is any change in the firm's production possibilities, the prices they expect to pay for the variable inputs, or in their objectives. If the prices of one or more of the variable inputs are reduced, then firms will plan to sell more at each price than before. This increase in supply is shown in Diagram 40 by a shift in the supply curve from S_1S_1 to S_2S_2. If there is no change in demand, then the price of this kind of cloth will tend to fall from *OP* to *OS* per yard. For any given change in supply, the ensuing change in price will be the greater the less is the price elasticity of demand, and it will be the less the greater the price elasticity of demand. The effects of simultaneous changes in demand and supply, whether in the same or opposite directions, can be illustrated simply on the same kind of diagram.

It must be emphasised that the preceding analysis only explains changes in the *relation* between the prices of cloth and the prices of other things. Thus, Diagram 39 shows us that if households' preferences for this cloth become stronger, its price will

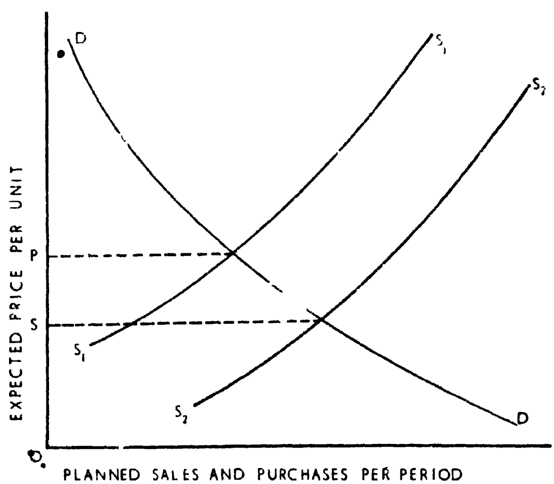


DIAGRAM 40

rise as compared with (*a*) the prices of other goods that they might buy; (*b*) their planned consumption expenditures which depend mainly on their incomes, which are merely the prices at which the households are currently selling the productive services that they own; (*c*) the prices of the variable productive services that are used to produce it. Similarly, Diagram 40 shows us that if there is a reduction in the prices of the variable inputs that are used by firms producing this cloth, then its price will fall as compared with (*a*) the prices of other products; (*b*) households' planned expenditures which depend ultimately on the prices at which they are selling their productive services; (*c*) the new and lower prices of the variable inputs.

We have so far shown that in the short-run the price of a commodity will be at the level at which the planned sales by existing firms will be equal in each period to the planned purchases of households, and that if there is an increase in demand, for example, price will rise to the level at which this condition is again fulfilled. We have not yet attempted to explain how, or by whom, the price is driven up, nor have we described the precise path by which it moves from the initial to the new equilibrium position. Initially, we shall suppose that the movement to the new equilibrium price is effected by a single intermediary (or group of intermediaries acting in concert), who works without either thought or expectation of reward, so that the price at which he buys is that at which he sells. This provides a pedagogically useful model of the adjustment of price towards its equilibrium level, though it is difficult to find any actual markets in the real world to which it is a close approximation. We shall assume that the production period for firms is the same as the purchase period for households, each being equal to one week; that sales and purchase plans are made at the beginning of the week on the basis of the price that is expected to rule during it, and that once made these plans are unalterable until the beginning of the next week. Let us now suppose that there is a permanent increase in demand at the beginning of week 1: that is, that the demand curve in Diagram 41 moves from D_1D_1 to D_2D_2. If firms and households have already laid their plans on the expectation that the price *OP* will rule, then during week 1 firms will supply the intermediary or merchant with *OH* yards to sell, and he will become aware (through orders that he is un-

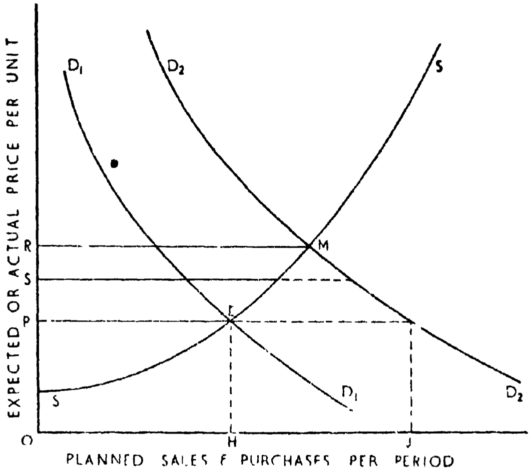


DIAGRAM 41

able to satisfy) that this falls short of the amount that households want to buy at OP per yard. For week 2, therefore, the merchant will plan to buy more from firms, but to induce them to produce more a higher price—say, OS—must be offered. If the price is fixed at OS for week 2, the merchant will find that his experiences of the first week are repeated though in lesser degree. He will plan a further increase in his purchases from firms for week 3, and these adjustments will continue until the price has reached OR per yard, for only then will the flow of cloth in each week from firms to the merchant be just equal to the flow of cloth from him to the households. The description of what would happen on these assumptions if there had been a reduction in demand is similar and it is left to the reader. In this model, the price, in moving to the new equilibrium level, follows the path traced by the short-run supply curve between L and M.

We may alternatively assume that the product is a perishable one, so that it must all be sold within the period in which it is produced. If we again suppose that the firms producing it expect the price OP to obtain in week 1, they will plan to produce OQ. If there is a spontaneous increase in demand at the beginning of week 1 from D_1D_1 to D_2D_2 in Diagram 42, then the price in that week will rise to OA. This increase in price may be effected by wholesalers or merchants, who are more or less aware of the enhanced demand, and who, desiring to maximise their net

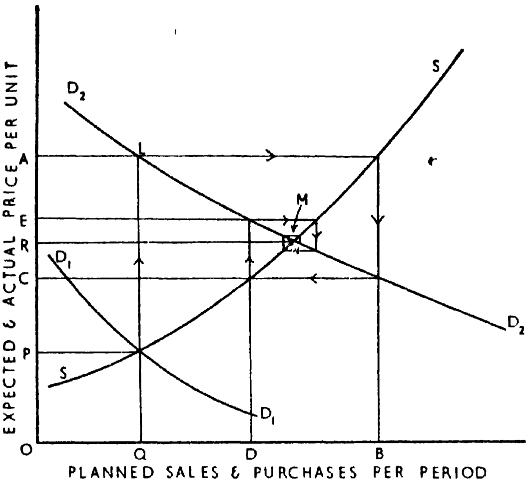


DIAGRAM 42

revenues, buy 'cheap' and sell 'dear' to households. Or it might be the result of those households that were first in the queue acquiring *OQ* from firms at *OP* per unit, and re-selling to those behind them in the queue, these transactions continuing until the price was such that no household possessing the commodity was willing to re-sell and no household wanting it willing to buy — that is, until the price had reached *OA* per unit. We may call *OA* the *market* equilibrium price, to distinguish it from the short-run equilibrium price like *OP* or *OR*. What will happen to the price in the weeks that follow will depend mainly on how firms revise their production and sales plans. We shall explore briefly what would happen if each firm always expected the price in the period lying ahead to be that which ruled in the present period.

If each firm expects the price *OA* to obtain in week 2, then together they will plan to produce a quantity *OB* in week 2, for in the light of their price-expectations that quantity alone will promise to maximise their net revenues. When the quantity *OB* is actually offered for sale, the price per unit will fall to *OC*. If each firm expects the price to be *OC* in week 3, they will plan to produce *OD* — in week 3, then, the price must rise to *OE*. We can see from the diagram that, on these assumptions, the price will gradually approach the new equilibrium level, *OR*. The path by which the price moves from *OP* to *OR* can be seen more

clearly from Diagram 42(*a*), where we measure time (in 'weeks') on the horizontal axis, and the price that actually ruled in each week on the vertical axis. In this diagram, price is represented by '*p*' and the suffixes 0, 1, 2, ... ∞ denote weeks.

The fact that price fluctuates, rather than rises monotonically, towards the new equilibrium level is a necessary consequence of our assumption about the basis of the price-expectations of firms. The fact that we have a convergent fluctuation in Diagram 42 is because the new demand curve, D_2D_2, has a smaller slope at each price than the supply curve. If demand had had a greater slope than supply at each price, we would have had divergent fluctuations; if the two curves had had the same slope at each price, there would have been continuous fluctuations. These two possibilities are illustrated in Diagrams 43 and 44 respectively.

These consequences of our assumption that each firm expects this period's price to rule in the next period are called the Cobweb Theorem, because of the appearance of the diagram on which they are illustrated. Even if the other circumstances are favourable — a perishable commodity, no single producer of which can affect its price — it is unlikely, however, that a 'cobweb' fluctuation will develop: sooner or later entrepreneurs must observe that the assumption on which they base their price-expectations is being proved wrong by events, and that periods of high and low prices alternate with one another. Once this is realised the cobweb fluctuations will be neutralised, for

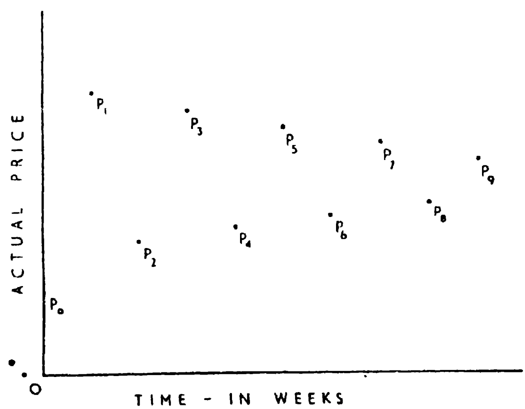


DIAGRAM 42 (*a*)

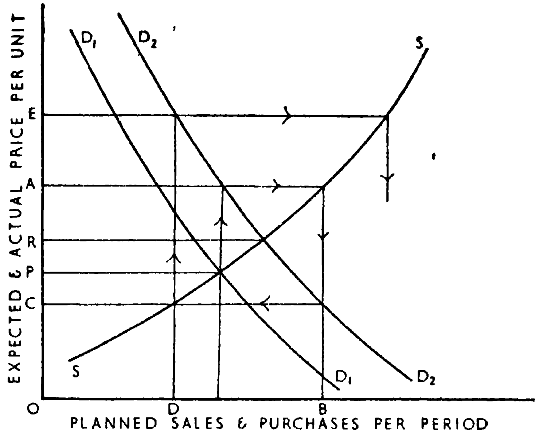


DIAGRAM 43

the more perspicacious firms will expect price to be low in the next period if it was high in this period (and vice versa), and make their production and sales plans accordingly. If the demand for the product rises, driving its price up to *OA* in Diagram 42, the price will probably fall monotonically in the ensuing periods, following the path traced by the range *LM* of the new demand curve. This sharp rise in the price of a commodity, followed by a continuing decline to somewhere above its initial level, is a not infrequent consequence of actual increases in demand. In practice, it is explained in part by the manner in

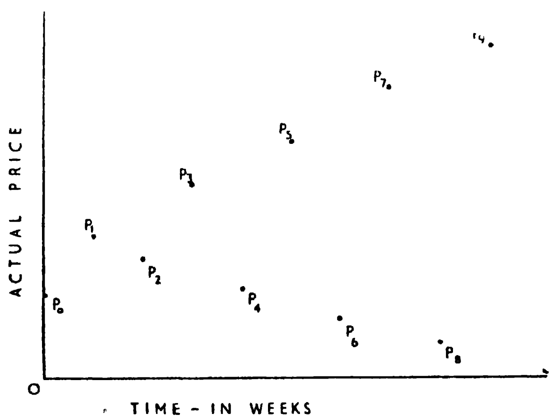


DIAGRAM 43 (*a*)

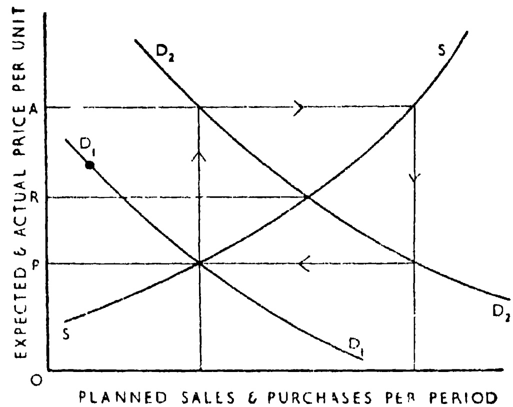


DIAGRAM 44

which firms revise their price-expectations (and it is on this that we have concentrated in our analysis above); it is in part due also to the fact that not all firms can employ more of the variable inputs — that is, can 'move along' their short-run supply curves — with equal ease. Those that are favourably placed can increase production quickly, but some time may elapse before others do so. Consequently, even if each firm knew what the equilibrium price was going to be, the quantity supplied would increase only gradually from period to period, causing the price to follow some path like *I M*.

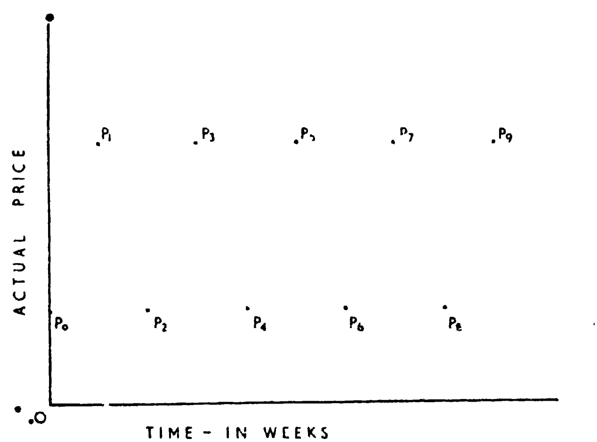


DIAGRAM 44 (*a*)

SHORT-RUN DEMAND AND SUPPLY ANALYSIS: ITS USES

The demand for, and the short-run supply of, a commodity explain the level towards which its price will tend while the firms' activities are circumscribed by past commitments. While the price is assumed to lie beyond the control of any single buyer or seller, the ultimate effect of all buyers and all sellers revising their expectations of what the price will be, and adjusting their planned purchases and planned sales accordingly, is the emergence of a situation in which all their expectations and plans are fulfilled. This explanation of the determination of relative prices in terms of demand and supply analysis has two main uses: first, it provides a number of headings under which we may conveniently classify the causes of changes in relative prices; second, it helps us to predict the consequences of price controls, of taxes on commodities, and of other similar measures.

Let us suppose that during the past month the price of eggs has risen as compared with the prices of all other things. A knowledge of elementary demand and supply analysis enables us to organise our quest for the cause of this event. If the relative price per dozen eggs has risen, the explanation must lie in changes in demand or in supply or in both of these. We have called the kinds of reason why demand or supply might change the 'determinants' or 'conditions' of demand and supply, and these were last listed on page 104 above: these provide us with a broad classification of the possible mediate causes of changes in relative prices. The next step is to discover which of these were operative. The system of classification that demand and supply analysis provides is, then, an aid to diagnosis: we observe the symptom which is a rise in the relative price of eggs, and the analysis tells us on what kinds of change we should focus our attention.

The usefulness of a classification may be illustrated by a few examples from other disciplines. The Morris Minor (Series II) Operation Manual lists* the mediate causes why the engine will not start: the catalogue runs from the ignition not being switched on, or the petrol supply being exhausted, through fouled sparking plugs, to a dead battery. By checking through this list, the cause can be discovered. Again, Dr. John Gibbens in *The Care*

* Third edition, page 25.

of Young Babies lists* the mediate causes of why babies cry: they may be hungry, thirsty, wet, dirty, hot, sweaty, lonely, neglected, over-stimulated, uncomfortable, or in pain. By elimination, the cause can be found, or, at least, the range of uncertainty may be narrowed.

None of these catalogues, of course, gives the *ultimate* cause of the event that we are seeking to explain. Thus, the engine may not start because the sparking plugs are fouled, or the baby might be crying because it is in pain. But these merely pose new problems rather than provide solutions, for we must in turn try to discover why the plugs are oiling up or what is causing the pain. We may discover that the relative price of eggs has risen because the preferences of households for them have become more intense; this in turn might be the result of any one of an almost innumerable list of causes, ranging from climate to caprice, and to explore further we need another classificatory system. Or the mediate cause might seem to lie in a rise in the prices of the variable inputs used by the egg producers: to probe further we can here use the classifications implicit in demand and supply analysis. The value of a classification must be judged by the help it gives in unravelling the problem at hand, and on this test all the classifications that we have listed are tolerably good. They help us to localise the mediate causes of the event in which we are interested, and if we wish to probe further they suggest where we should look for more information. If the plugs are oiling up, ask a skilled mechanic; if the baby is in pain, take him to a doctor; if tastes and preferences have altered, consult a social psychologist or anthropologist.

We have so far used demand and supply analysis to work from an event to its proximate cause, and we have seen that it undoubtedly clarifies hindsight. The analysis may also be used to deduce from an event its probable consequences. If we observe, for example, that the price of poultry feed has risen, or that the government has controlled the price of eggs or imposed a tax on them, we may, assuming *cetera paria*, predict what will happen to the price of eggs. Whether or not our predictions are proved by events to have been correct will depend on whether the *cetera* have indeed remained *paria*. Good economic predictions can never result from the mere mechanical application of the analy-

* Third edition, pages 175–6.

sis, for in a developed economy most consequences have several causes. Before hazarding a prediction in practice, we must decide whether any of the other determinants of demand and supply are likely to change. The analysis tells us what things to look at, but our judgement of how they are likely to alter and of how we should weight the probable changes in different determinants before chancing a prediction of their net effect is more a matter of 'feel' — that is, of that rather rare ability to measure the incommensurable and add the non-additive. Bearing this in mind, we shall explore in a more or less mechanical way the probable effects of price controls and taxes on commodities.

First, price controls. Let us suppose that there is an increase in demand and that the government makes it illegal for sellers to raise the price above its initial equilibrium level. At the legal maximum price (*OP* in Diagram 45), firms will plan to supply *OH* per period and this falls short by *HJ* of the quantity that households are planning to buy. If the price control is effective, this situation can continue indefinitely, for the households, or the merchants who serve them, dare not offer the higher prices that would alone eliminate this 'excess demand'. If price is thus prevented from distributing the quantity that is supplied of the commodity amongst all those who demand it, other methods must be found. Sellers may allocate the quantity *OH* amongst those desiring *OJ* on the basis of 'first come, first served', or they

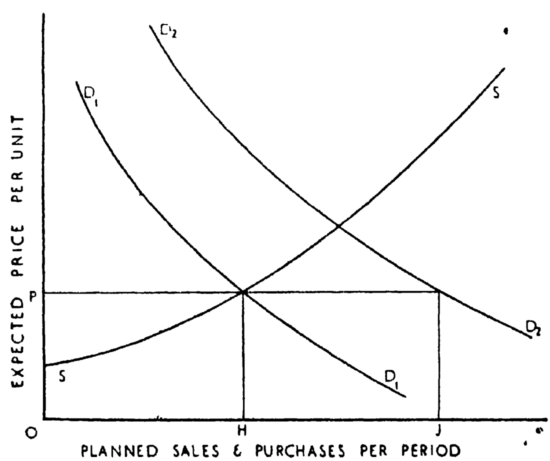


DIAGRAM 45

might hoard it 'under the counter' and distribute it on the basis of their personal feelings towards their customers or of their customers' past purchases and behaviour. When there are maximum price controls, some form of rationing is necessary, but it may be deemed socially undesirable that the choice of method should be left to sellers. In these circumstances, the government may issue to each household ration coupons, the value of each coupon being so fixed that all together they can 'buy' only the quantity of the commodity that is being supplied in each period.

It may be that at the equilibrium price (*OP* in Diagram 46), the total revenue that most firms earn per period is not sufficient to cover their total costs of production, so that sooner or later the number of firms producing the commodity will be depleted by bankruptcy. There may be political, social, strategic or humanitarian reasons why this is deemed undesirable by the government. To prevent it, the price may be fixed or 'pegged' above its equilibrium level — say, at *OS*. At the legal minimum price *OS*, the firms will plan to produce *OM* in each period, and this will exceed the planned purchases of households by *KM*. The individual firms, whose financial straits led the government to fix the price at *OS* per unit, will not be able to accumulate stocks at a rate of *KM* per period: if left to themselves, they will offer to sell at lower prices in an effort to dispose of their

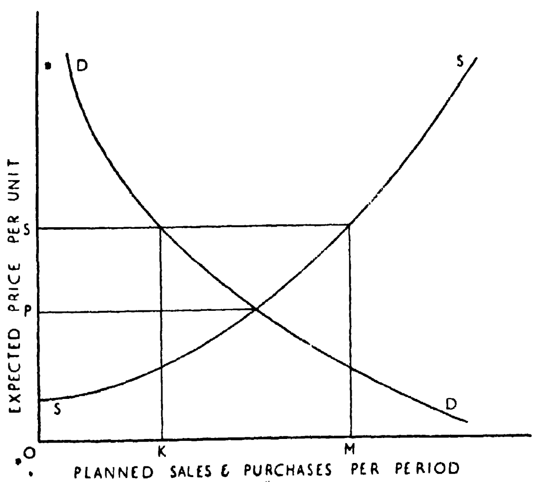


DIAGRAM 46

total output in each period and the price will fall to *OP*. To avoid this, the firms may be required to restrict their planned outputs in each period to *OK*. Or the government may set up a central agency to buy *KM* in each period at the legal minimum price, and destroy it (as happened with Brazilian coffee in the inter-war years), or store it (as now happens in the United States with many agricultural commodities). If the latter, it may be hoped that the demand for the commodity will rise (or the supply of it will fall) in the future, so that the price *OS* will fall short of the then equilibrium level; if this should happen, the accumulated stocks could then be run down to meet the 'excess demand' for the product at the legal minimum price.

Next, the effects of taxes. Let us suppose that before the imposition of the tax the conditions of demand and supply are as illustrated by D_1D_1 and S_1S_1 curves respectively in Diagram 47, and that the equilibrium price is *OP* and sales and purchases per period each equal to *OQ*. Let us now suppose that the government decides to exact a fixed sum (say, threepence) from each seller for each unit that he sells — that is, that a specific tax is imposed. The immediate effect of the tax will be to shift the supply curve due northwards through a distance equal to the tax per unit. For each seller will argue as follows: given my costs of production, I shall only plan to produce (say) 100 units

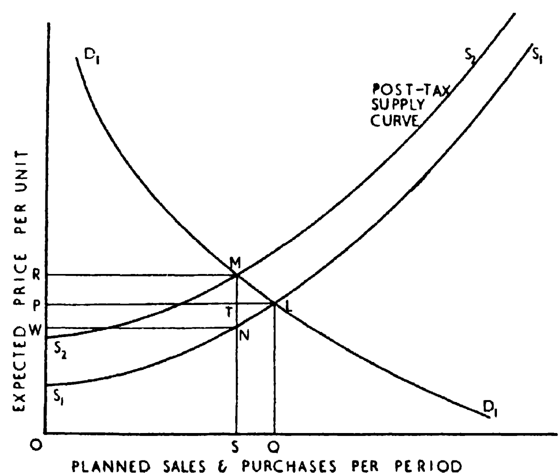


Diagram 47

if I expect to receive 1s. per unit; now that the government exacts 3d. from me for each unit that I sell, buyers must pay me 1s. 3d. per unit if I am to continue producing 100 units in each period, for it is only if they do so that I shall be left with the 1s. per unit that I must get if my net revenue is to be at a maximum at this output. After the imposition of the tax, then, the equilibrium price of the commodity will rise to OR and the planned sales and purchases in each period will fall to OS. By comparing the initial equilibrium (L) with the post-tax equilibrium (M) in Diagram 47, we can measure the effects of the tax. Buyers now pay PR ($=MT$) per unit more for QS ($=LT$) less of the commodity; sellers now receive TN less per unit for their lower sales of OS in each period. In the post-tax equilibrium, buyers expend $OSMR$ per period; of this sum sellers pass on $WNMR$ to the government and they are left with $OSNW$. In terms of the diagram, of the tax of MN per unit that has been imposed, we may say that MT is 'paid' by buyers and TN 'paid' by sellers.

When the tax is a relatively small proportion of the price of the commodity, it can be shown that the ratio of MT to TN is equal to the ratio of the elasticity of supply (in the range LN of the supply curve) to the price elasticity of demand (in the range LM of the demand curve).* If we know these relevant elasticities, we could predict the relative impact of the tax on price and output: the less elastic is demand and supply, the less will output fall and the more will price rise; the more elastic is demand and supply, the more will output fall and the less will price rise.†

* In Diagram 47, $E_s = \frac{SQ}{OQ} \Big/ \frac{WP}{OP}$, and $E_d = \frac{SQ}{OQ} \Big/ \frac{PR}{OP}$.

Therefore, $E_s/E_d = MT/TN$.

† Let t be the tax per unit: then $t = MT + TN$.

Now, for MT we can write $E_s/E_d \,.\, TN$.

We then get: $$t = \frac{E_s}{E_d} . TN + TN = TN\left(\frac{E_s + E_d}{E_d}\right).$$

Hence: $$TN = t\left(\frac{E_d}{E_s + E_d}\right).$$

Similarly: $$MT = t\left(\frac{E_s}{E_s + E_d}\right).$$

If $E_d = 0$, then $TN = 0$, and $MT = t$.
If $E_s = 0$, then $TN = t$, and $MT = 0$.
If $E_s =$ infinity, then $TN = 0$, and $MT = t$.
If $E_d =$ infinity, then $TN = t$, and $MT = 0$.

In presenting demand and supply analysis and in examining some of its more obvious applications, we have tacitly assumed that in each period planned production was identical with planned sales, and planned purchases the same as planned consumption. Consequently, in our diagrams, the flows of production, sales, purchases and consumption were equal to one another in each period. If the commodity can be stored cheaply, this assumption is untenable, and when we drop it, other patterns of price adjustment than those we have already examined become possible. By way of example, let us suppose that there is a rise in the demand for some commodity and that its price starts to rise. We shall suppose also that this creates expectations in the minds of buyers and sellers that the price is going to continue to rise in the future. Since the commodity can be stored cheaply these expectations will cause a change in the distribution over time of purchases and sales. Buyers may still desire an even flow of consumption in each period, but they will plan to concentrate their purchases in the present when the commodity is relatively cheap at the expense of the future when they believe that it will be relatively dear. Purchases will exceed consumption now so that buyers accumulate stocks; planned purchases will fall short of expected consumption in the future when these stocks are being depleted. Sellers will plan to maintain or increase production and reduce sales now; in this way, they will build up stocks that can be used to supplement current production in the future, when (or if) their expectations are fulfilled. Thus, if price is expected to rise in future, the increase in present purchases and the curtailment of present sales will tend to raise the price more quickly now and so justify these expectations; the reduction in purchases and the increase in sales in later periods will arrest the rate of increase in prices — indeed, it may even cause the price to begin falling and so create expectations of further price reductions.

If all this was initiated by a permanent rise in demand, it is probable that the price will ultimately settle at its new and higher equilibrium level. The path by which the price reaches this level may be similar to that shown in Diagram 42(*a*). Such paths are the more likely the less aware are buyers and sellers of the nature and strength of the true cause of the initial rise in price, namely, in our example, a permanent rise in demand.

PRICE DETERMINATION: LONG-RUN

In this chapter so far we have confined our attention to the short-run. If the demand for a commodity rises (or falls), the firms who are producing it are limited in their responses by the possession of fixed factors. As the past commitments that fix the quantities of certain factors at their disposal fall due for renewal, firms will be able to make a more complete adjustment to the new demand conditions. In the remainder of this chapter, we shall examine the nature of these adjustments and describe how they are likely to affect relative prices.

In the short-run, the output of the product can only be varied within the limits set by the fixed plants of existing firms. In the long-run, the output of the product can be varied by the firms who are already producing it increasing or reducing their scale of operations by changing the nature and size of their 'fixed' factors; it can be varied also by new firms entering the industry, or by firms that are already there retiring. We have seen in Chapter 3 that the range of choice facing an entrepreneur who is planning to enter or remain in an industry is summarised in his planning curve. Each entrepreneur will have a planning curve for each industry that he might enter. Given the price at which he expects to be able to sell each product, and his knowledge of the methods by which it might be produced and his expectations of the prices of the inputs these require, he can deduce the maximum net revenue per period he could get were he to decide to make it. He will decide to produce that product — that is, to enter that industry — that promises him the *maximum maximorum* of net revenue. The same choice will face an entrepreneur who has rid himself of all past commitments in an industry, for he can choose whether he will remain there or set up in another industry.

Let us now suppose that the demand for, and short-run supply of, some product are as illustrated by the D_1D_1 and S_1S_1 curves in Diagram 48, and that at the price OP and output per period of OQ there is long-run equilibrium — that is, were the price to continue at this level, no new firm would desire to enter this industry and no existing firm would plan to leave it or vary the scale of its operations in any way. We shall presently describe how a position of long-run equilibrium may be reached. Let us

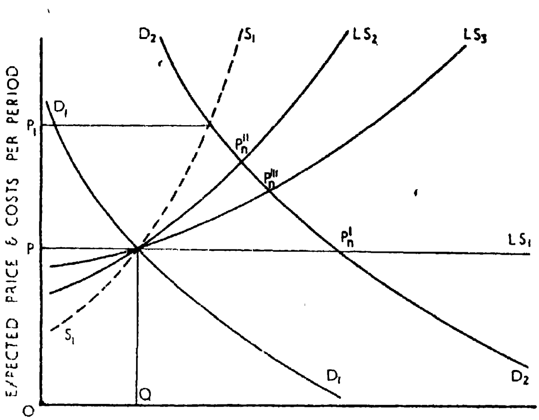


DIAGRAM 48

suppose that there is a sudden but permanent rise in the demand for this product to D_2D_2, and that the demand is generally expected to remain indefinitely at its new level. We shall assume also that no change is expected to occur in the conditions of demand and supply in any other industry. In the short period, in which no change in either the number or size of firms is possible, the price of the product will tend towards OP_1. The level towards which the price will tend in the long-run will depend on the elasticity of the long-run supply curve. The path by which the price will move towards its long-run equilibrium level will depend on the expectations that each entrepreneur has about the price of the product and the prices of the inputs needed to produce it, when he is making his long-run decision.

We have shown (page 101 above) that the long-run supply curve of a firm will be perfectly elastic until the minimum point of its planning curve is reached and that it will then rise, the rising portion coinciding with its long-run marginal cost curve. Given the prices of the relevant inputs, the elasticity of the rising part of the firm's long-run supply curve will depend on (*a*) the physical production possibilities, and (*b*) the 'diseconomies' of large-scale management — that is, the probability that as the rate of output is increased, the problem of co-ordinating the activities of the greater variety and quantity of inputs that are required will tend to raise unit costs of production. The elasticity of the long-run supply curve of the product depends not

only on these but also on the relationship between the minimum points of the planning curves of the different entrepreneurs who might enter the industry. If there are a very large number of entrepreneurs who might remain, or start work, in this industry in the long-run; if they all have the same expectations about the prices of inputs and if the minimum expected net revenue required to induce each to enter this industry is the same, and if they are all equally knowledgeable about production possibilities and equally competent as co-ordinators, then their planning curves will all be identical with one another. When their long-run supply curves are added together, the resulting long-run supply curve of the product will be perfectly elastic,* and, as we can see from Diagram 48, the price in the long-run will return to its initial equilibrium level, *OP*.

If the actual or potential entrepreneurs in the industry are not equally knowledgeable about production possibilities, or if they differ from one another in the ability to make decisions and determine policy, or if the minimum expected net revenue needed to induce each to enter the industry is not the same for all of them, then each will have a different planning curve. The planning curves may differ in that their minimum points come at different levels or at different outputs. The minimum point of an entrepreneur's planning curve may be at a relatively high price because (*a*) he is unaware of some methods by which the product might be produced, or (*b*) he is less able than some of his fellows to co-ordinate effectively, or (*c*) he requires a relatively large expected net revenue to attract him to this industry. The minimum point of an entrepreneur's planning curve is roughly explained, therefore, by his relative 'efficiency' in the industry to which the curve relates and by his relative 'efficiency' in the other activities in which he might indulge: the less is the former and the greater the latter, the higher will it be, and vice versa. In these circumstances, the long-run industry supply curve will be less than perfectly elastic, and the level towards which the price of the product will tend in the long-run will be between *OP* and OP_1. In general, we may say that the supply curve will be the more elastic, and the price will ultimately be nearer to *OP*, the smaller are the differences between the minimum points of

* The rising part of the firm long-run supply curves will only begin affecting the shape of the long-run industry supply curve at an infinitely large rate of output.

the individual planning curves; and the greater are these differences, the less elastic will be the supply curve, and the nearer will be the long-run equilibrium price to OP_1. These conclusions are illustrated by the curves LS_3 and LS_2 respectively in Diagram 48.

The points P'_n, P''_n, and P'''_n represent long-run equilibria, for at each of these prices (on the appropriate assumption about the elasticity of supply) the quantity of the product that households plan to buy would be the same as the quantity that the firms plan to produce and sell in each period. At each of these prices, each firm's output would be at the level that promised it the maximum net revenue, and there would be no incentive for any new firm to enter the industry or for any existing firm to leave it.

The path that price follows when moving to its long-run equilibrium level will depend primarily on each entrepreneur's expectations about the prices at which he hopes to be able to sell his product and buy his inputs during the ensuing long period. The time taken for price to traverse this path will depend mainly on how quickly long-run decisions can be made in terms of calendar or clock time. Initially, we shall suppose that each entrepreneur expects the price at which the commodity is now being sold, and the prices at which inputs can now be bought, to rule indefinitely, and that long-run decisions can be implemented quickly.* If the industry is in long-run equilibrium at OP in Diagram 49, and if there is a rise in demand for the product to D_2D_2, the price will soon rise to OP_1, the short-run equilibrium. As each firm makes its long-run plan on the assumption that OP_1 will rule indefinitely, the planned output per period will rise to P_1M_1 when these plans have been implemented. This exceeds planned purchases by households at OP_1, so that price will tend to fall, and as it falls, each firm will contract its output along its short-run supply curve. The price will therefore fall to OP_2, where S_2S_2, the new short-run supply curve, cuts D_2D_2. If each firm again supposes that the price of the pro-

* A long-run decision may be implemented quickly if time is measured in days, hours or weeks: yet once implemented, the long-run decision may bind the firm for a long-period of calendar time. It is probable that this assumption is reasonably true of retail trading, and of many industries that supply personal services. It may also be true in agriculture: thus, a farmer may plan in autumn to devote all his land to growing oats in the ensuing crop-year, and this decision, once made, will bind him for the ensuing calendar year.

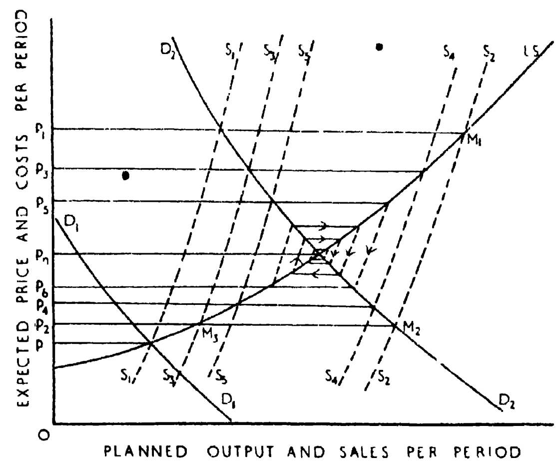


DIAGRAM 49

duct will remain at OP_2, together they will plan a long-run output of P_2M_3 per period. This will fall short of the planned purchases at the price OP_2, and price will tend to rise. As price rises, each firm will expand output along its short-run supply curve, so that the price will tend towards the level OP_3, where S_3S_3, the short-run supply curve of the industry when all firms have the fixed plant, etc., appropriate to the point M_3 on the long-run supply curve, cuts D_2D_2. And so on. The diagram shows that there will be a convergent fluctuation towards the long-run equilibrium price OP_n, and that, given the elasticities of the D_2D_2 and LS curves, the rate at which price converges on OP_n will be the greater the more elastic are the short-run supply curves.*

We would not expect to find the price of a product actually fluctuating in this way, and for three reasons. First, as many long-runs as there are 'turns' in the 'spiral' must elapse before price reaches its long-run equilibrium level, and during this time it is likely that there will be changes in demand, techniques, input prices and the alternative opportunities open to entre-

* If the short-run supply curves in Diagram 49 were all perfectly inelastic, Diagram 49 would be identical with Diagrams 42, 43 and 44. In the long-run we are unlikely to get either a divergent or a continuous fluctuation because (*a*) the long-run supply curve will probably be fairly elastic; (*b*) even if LS is less elastic at each price than D_2D_2, the short-run supply curves always have some elasticity, and their influence will probably overcome the tendency towards divergent or continuous fluctuations.

preneurs. Second, even if there are no changes in the conditions of demand and supply, each entrepreneur is likely to revise his belief that the price will remain at its present level. Each entrepreneur will know that, while the price of the product lies beyond his control, its level will depend on the total output of the product. In making a long-run plan, each entrepreneur may be aware that others have made, or may suspect that others may make, similar plans, and that the price of the product will probably fall in the future. In deciding upon what plant to build he may therefore assume that the price of the product, when the plant is in operation, will be somewhere below OP_1. Third, even if neither of these reasons is operative, the price might nevertheless move more or less directly towards OP_n, for long-run decisions are likely to be implemented *seriatim* rather than simultaneously. When the demand for the product rises, some existing firms whose past commitments are lapsing, or some firms new to the industry, may make long-run plans. As these plans are put into effect, the price of the product will start falling. By that time, other firms may find it possible to make their long-run decisions, and in doing so they will be influenced both by the behaviour of price since the rise in demand and by its existing level. As they make and implement their plans, there will be a further fall in price. In other words, it may take a very long time for the LS curve in Diagram 49 to become fully operative, and while it is becoming operative, the price of the product will be falling towards OP_n along the demand curve D_2D_2.

LONG-RUN DEMAND AND SUPPLY ANALYSIS: USES

Demand and supply analysis, in the long-run as in the short-run, has two main uses: first, it provides us with a number of headings under which we may meaningfully classify the causes of changes in relative prices; and second, it helps us (though to a very modest extent) to predict the ultimate effects on relative prices of present events, like a growing demand for the product or the discovery of a new method of producing it.

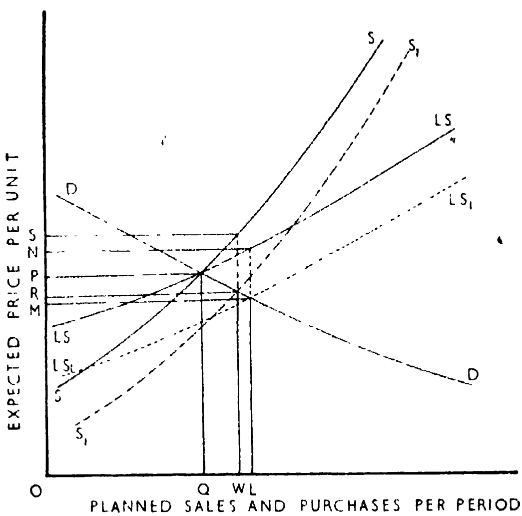
Let us suppose that during the past year or so, the price of eggs has continuously fallen as compared with the prices of other things. The analysis of this chapter tells us that the explanation must lie in changes in demand and supply. The price may have declined, for example, because demand has been falling,

with existing firms contracting their outputs along their short-run supply curves; or because demand rose some time ago, and the producers have been adjusting along their long-run supply curves; or it may be that the explanation lies in changes in the conditions of supply, which have caused rightward shifts in either the short-run or long-run supply curves, with demand conditions remaining unaltered. Our knowledge of the economic history of the past few years may tell us whether or not demand has altered, and we have seen earlier in this chapter how the analyses of Chapter 1 can help us to locate the proximate causes of the changes in demand that have occurred. If we feel that demand has not changed, and if, in addition, we observe that the number of firms and their size (as measured by the amount of 'fixed' factors each employs) have remained more or less the same, we can conclude that the fall in price is largely due to rightward shifts in the short-run supply curve; and we have already shown how the analyses of Chapter 2 can help us to discover why this may have happened. If demand has not changed, but if the number and/or size of the firms producing eggs have risen, then the explanation of the fall in price is probably to be found in a rightward shift in the long-run supply curve. The analysis of Chapter 3 provides us with a classification of the reasons why this might have happened: the determinants of long-run supply are (*a*) the methods or techniques of production; (*b*) the prices of the inputs that are required; (*c*) the firms' objectives. The discovery of a new method of production will lower the minimum point of each firm's planning curve and probably move it to the right, for the urge to seek new methods of production springs from the desire to reduce costs. A fall in the price of inputs will have the same general effect. It may be that the input whose 'price' has fallen is entrepreneurship; this would happen if there were a fall in the maximum net revenue that entrepreneurs could expect to earn in industries other than egg production. Thus, if for some reason other agricultural activities become less profitable, the 'price' that each entrepreneur would want for his services in egg production would fall.

Long-run demand and supply analysis is more useful in clarifying hindsight than in informing foresight. If the government decides to subsidise the production of some commodity, we can make a fairly firm prediction of the short-run consequences:

the short-run supply curve will fall vertically through a distance equal to the subsidy per unit as in Diagram 50; the price that households pay will fall from *OP* to *OR*; the price that sellers receive will rise from *OP* to *OS* (= *OR* plus the subsidy per unit of *RS*), and sales and purchases will rise from *OQ* to *OW*. Our prediction may be quite accurate, for the short-run, which we have defined in terms of operational time (that is, as time during which certain changes can or cannot take place), can usually be related to a short period of calendar time or clock time, and the shorter the calendar time the greater the likelihood that demand and supply will not alter. In forecasting the more immediate effects on price and output of a subsidy, we may then be justified in assuming that the new equilibrium is reached by firms and households 'moving' along their existing demand and supply curves.

If before the subsidy is introduced, the industry is in long-run equilibrium at the price *OP* and the output *OQ*, and if in the light of the existing conditions the long-run supply curve is *LS*,



SS = Short-run supply curve before subsidy
S_1S_1 = Short-run supply curve after subsidy
LS = Long-run supply curve before subsidy
LS_1 = Long-run supply curve after subsidy
RS = MN = Subsidy per unit

we may deduce (from a mechanical application of the analysis) that, if the subsidy continues, the price paid by households will ultimately fall to OM per unit and the output per period will rise to OL, with firms receiving ON ($= OM$ plus the subsidy MN) per unit. Events will only confirm this prediction if all other things remain unchanged while the adjustments in supply are taking place. The long-run adjustments, however, may require months or years, and in the meantime the determinants of demand and supply will almost certainly alter: long-run changes, for example, may be occurring in other industries and the changes in the prices of their products will affect the conditions of supply and demand for the commodity in which we are interested. It would be wrong to conclude, however, that long-run demand and supply analysis is valueless as an aid to prediction. In the first place, it helps us to deduce the ultimate effects on price of any once-for-all change in the conditions of demand and supply, *ceteris paribus*. 'The only complete or logical procedure' is to describe 'the action of any force by stating the *final* condition which it tends to bring about, the conditions under which it would cease to work'.* In the second place, while it seldom helps us to foresee what other changes will occur in the planning data of firms and households as the adjustments to some initial change are proceeding, it does provide an analytical framework within which these changes can be interpreted. We can, therefore, modify our initial prediction as events unfold.

However, not all changes in the determinants of long-run supply are equally unpredictable: there are some that we, as economists, might foresee, even though no firm in the industry may take account of them when laying its long-run plans. Economic history suggests that most industries that have expanded have at least two things in common. First, the development of new methods of production, and of new variants of the product, usually proceeds *pari passu* with the expansion of the

industry; and second, that as the industry expands, the prices of some of its inputs tend to rise while the prices of other inputs tend to fall.

During the last two centuries, the rate of technological development has been increasing more or less continuously in most industries. In most developed economies, there is probably a general expectation that the flow of new techniques and of new products will continue in the future, and this is almost as strong as the general expectation that the flow of foodstuffs from the farms or of textiles from the factories will continue. This expectation would not by itself destroy the analytical and prognostic value of the concept of the long-run supply curve, for if all new developments were 'acts of God' or if they appeared from outside the industry, the long-run supply curve would remain as a useful summary of entrepreneurs' intentions. Each entrepreneur would make his plans on the basis of present techniques, and he would revise them as best he could as and when these random or stochastic developments occurred. In practice, however, technological research is increasingly being carried out by firms, and the 'production' of new techniques and of new knowledge is planned in the same way and on the same principles as is the production of more tangible products. The planning horizon for new methods and new products may be longer than that for long-run plans of the kind we have already described. And the choice of a 'research' plan differs more in degree than in kind from the choice of a long-run production plan as described in Chapter 3: the 'outputs' are less tangible and less predictable, but their 'volume' varies roughly with the quantities and the kinds of inputs (research workers and specialised equipment) that are used. We may then think of the actual long-run plan that an entrepreneur makes (especially if he is engaged in newer industries like plastics, radio, television, motor-cars, etc.) as being the compromise, which he considers potentially the most profitable, between the exploitation of existing methods and products and the quest for new methods and products.

We have not space to pursue this matter further: enough has been said to show that in many industries entrepreneurs will seldom devote all their energies to 'moving along' their long-run supply curves as we have defined them; some part of their re-

sources will be devoted to shifting these curves. Our long-run supply curves assume that all other things remain equal, but entrepreneurs will devote some effort to making some of them unequal. Conceptually, there is a choice: first, we may treat each firm as being a multi-product firm in the long-run, planning to produce both its existing product by existing methods and new methods or new variants; or second, we may view research as being the quest for a new input, the demand for which could be described in the same way as we shall later describe the demand for a machine or a factory building; or third, maintain our definition of the long-run supply curve and assume that technological progress is continuously tending to shift it to the right. We shall choose the last of these, for the present stage of knowledge about the causes of technical progress makes it almost impossible to posit the functional relationships that either of the other two would require. The long-run supply curve, then, shows us the long-run plan that an entrepreneur would make at a point in time at each price at which he expects to be able to sell his product; technical progress means that at each successive point in time, the long-run supply curve will, *ceteris paribus*, be further to the right.

As an industry expands, not only may new techniques appear, but the price that each firm must pay for some inputs may rise, while the price it has to pay for other inputs falls. Most industries begin with a few firms, producing a relatively small output. They may each produce the specialised equipment they require or have it made to order by firms in other industries. In either case, the relatively small demand for the machine confines the firm that makes it to the production possibilities that lie at the western end of the relevant planning curve, so that the cost per machine will be relatively high, and its price will be relatively high also. As the industry grows, the demand for the specialised equipment and services that it requires will rise also: the firm (or firms) that makes the machines, for example, can then 'move along' its planning curve (assuming no change in techniques, input prices, etc.); the cost per machine will fall as the number produced rises, and hence its relative price will tend to fall also. Reductions in input prices that come about in this way are called *external economies*; they are so-called because they are external to each firm that demands these inputs. If industry *A* ex-

pands, and if as a consequence the price of one of its inputs that is produced by industry *B* falls, this is an external economy for each firm in industry *A*, but it is the result of economies that are internal to each firm in industry *B*, for the rise in the demand for its product makes it profitable for each firm there to 'move along' its planning curve towards its minimum point.*

Examples of external economies spring easily to mind. As an industry expands, it may become profitable for new firms to specialise in collecting and disseminating market information, or in marketing the industry's product, or in supplying it with consultant services. If the expanding industry is localised geographically, the external economies may be more striking: its skilled labour may be trained at local technical schools (for at each school there may be enough pupils to make the employment of a full-time teacher worthwhile), and the public utility industries may evolve with it, being continuously adapted to its needs. In general, the external economies will be the greater the less are the differences between the products of the firms that enjoy them, and the more standardised are the inputs that are being demanded from other industries.

If the demand for industry *B*'s product is initially large enough to enable each firm to produce at an output that lies at or beyond the minimum point of its planning curve, then when industry *A* expands, the price that it must pay for the input it buys from *B* will, *ceteris paribus*, remain the same if *B*'s long-run supply curve is perfectly elastic, and rise if it has any degree of inelasticity. The tendency for the price of *B*'s product to rise may, as we saw earlier, be offset by the development of new techniques of production. As industry *A* expands, the price that each firm must pay for the labour-service it uses, may rise also, and the extent of the rise in price will depend, as we shall see later

* In the situation described above, we cannot assume, as we have been doing, that no single firm can affect the price at which its product is being sold. This assumption will only be valid (see *infra*, Chapter 9) if, *inter alia*, there is a very large number of firms producing the same product. This clearly cannot be the case initially in the machine industry in the above example before the rise in the demand for its product; for if there had been a large number of firms then, there would already have been a strong incentive for each to expand output and so reduce cost. Initially, then, the quantity of industry *B*'s product demanded per period by industry *A* must have been less than the cost-minimising output for a firm in *B*, with the existing techniques, etc. It may be, of course, that the expansion in *A* is so great that the quantity of *B*'s output demanded per period is large enough to support a large number of firms, each enjoying its internal economies.

(*infra*, Chapter 7) on the elasticity of supply of each kind of labour-service to industry *A*. The labour costs per unit of output in industry *A* may rise, not only because each firm must pay a higher price per unit for labour-service of the same quality, but because, while the price remains the same, the quality of the labour-service that can be hired falls. Increases in price that occur for these reasons are called *external diseconomies*. They are so-called because they are external to each firm in industry *A*: the rise in the price of the input is not caused by the expansion of any single firm in *A* but is rather the consequence of the expansion of the whole industry.

Analytically, the problems introduced by external economies and diseconomies are of the same order as those discussed on pages 124–6 above. There, we saw that if each firm planned for the long-run on the assumption that the price of the product would remain at its present level indefinitely, the ultimate effect of all firms implementing their plans would be to reduce the price of the product. Similarly, if each firm plans on the assumption that the price of each input will remain the same over the long-run, and if there are external economies or diseconomies, the ultimate effect of all firms implementing their plans will be a change in input prices. In both these cases, when the firms have put their long-run plans into effect, the actual net revenue which they will be earning in each period will differ both from that which they had expected to earn before their plans were implemented, and from the maximum net revenue they now feel they could earn were they to plan anew on the basis of existing input and product prices. There will ensue a period of adjustment and readjustment (like that illustrated in Diagram 49) culminating, after many long-periods have elapsed, in a new equilibrium in which all their expectations are being fulfilled.

SHORT-RUN AND LONG-RUN DEMAND

Thus far, in analysing changes in relative prices, we have assumed sudden and permanent changes in demand, and sought to discover the pattern of supply adjustments over time, and its consequences. We have split time into three operational periods: first, the market period in which no revision whatsoever can be made in the output plan in response to changes in the firm's ex-

pectations of the selling price of the product; second, the short-run, during which the output plan can be revised within the limits imposed by past commitments whose tangible embodiments are fixed factors; and third, the long-run, for which virtually no inputs are fixed, so that the fullest possible adjustment can be planned to changes in the expected selling price of the product. We have seen that if demand rises, price will rise to its highest level in the market period, that it will fall somewhat in the short-period, and that it will fall still further in the long-period, the *cetera* appropriate to each operational period remaining *paria*.

We have not so far sought patterns in the demand adjustments that occur as time passes. We shall do this now very briefly, for both the concepts and the analysis are analogous to those used for supply. We may define the market period for households in one of two ways. First, we may suppose that during the market period the quantity of each good that the household plans to buy cannot be altered: on this definition, the household's demand for each good that it plans to buy will be perfectly inelastic, so that if, on balance, prices are higher than the household expected them to be when laying its plans, it will spend more than it intended during the market period, and save less, and vice versa. Second, we may assume that during the market period the planned expenditure on each good is unalterable: on this definition, the household's demand for each good will have unit elasticity, and if in the market period the actual prices are different from the prices the household expected to have to pay when making its plans, the household will enjoy a lower satisfaction or ophelimity than it had hoped.

The household, like the firm, may have past commitments that limit the extent to which its purchase plan can be revised while they bind it. The members of the household may have insured their lives and contracted to pay the premiums in quarterly instalments, or they may be buying a television set or a motor-car on hire purchase and paying a fixed sum each month to the seller. In either case, the household will have certain fixed expenditures (the analogue of the firm's fixed costs) in each period. If the relative prices of one or more of the goods that it buys should change, the household will be limited to the combinations of goods that can be bought with its planned con-

sumption expenditure less its fixed expenses. Some adjustment in either the planned purchases of each 'variable' good or in the planned expenditure on it (or in both) will now be practicable, so that we would expect the household's demand curves for these goods to be rather more elastic in this 'short-run' than they were in the market period. When the household has rid itself of all past commitments like hire purchase agreements it may plan a full adjustment of its purchase plan to the new pattern of relative prices. For most goods and services, we would normally expect the household's long-run demand curve to be more elastic than its short-run demand curve.*

The influence of demand on the behaviour of prices over time can be illustrated on a diagram similar to Diagram 48. In Diagram 51, *LD* and *LS* are the long-run demand and supply curves respectively, and we shall suppose that initially the price is at the long-run equilibrium level *OP*. Let us now suppose that there is a sudden and permanent fall in the long-run supply to LS_1. If D_m is the 'market-period' demand curve, the immediate consequence will be a rise in price to OP_1. As households make their short period adjustments, the short period demand curve D_s will become effective, and the price will fall to OP_2. As their long period demand becomes operative — that is, as the *LD*-curve becomes the locus of alternative purchase plans open to the households — the price will fall further to OP_3. The use of the concept of long-run demand to predict the ultimate consequences of present events is subject to the same limitations as

* The analogy between long-run demand and long-run supply decisions might be pressed further, though to do so would provide more analytical excitement than insight. We might, for example, classify the consumption possibilities that are open to the household when making a long-run decision into 'standards of living' or 'methods' of consumption. The goods and services that the household might buy (its inputs) will vary both in size (partially true of houses) and in kind (bicycle and motor-car, ice and a refrigerator, coal fires and oil-fired central heating), from one 'standard of living' or 'technique of consumption' to another. Each standard of living will have its own set of 'fixed' inputs, like a house, a car, club-membership, school-fees, etc., and these will probably bulk the larger the 'higher' is the standard. The level of satisfaction that the household would enjoy can be varied within each standard by varying the quantities of the 'variable' inputs (current consumption goods) that are combined with the appropriate 'fixed' inputs. Given the expected prices of all inputs, we could calculate the level and behaviour of the planned total expenditure (both 'fixed' and 'variable') per period within each standard of living. Given the household's expectations of its planned expenditure in each period and of how it expects this to behave, it will choose that standard of living that promises the maximum satisfaction per period. If the household's satisfaction, like the firm's revenue, could be cardinally measured, the analogy might be pressed even further.

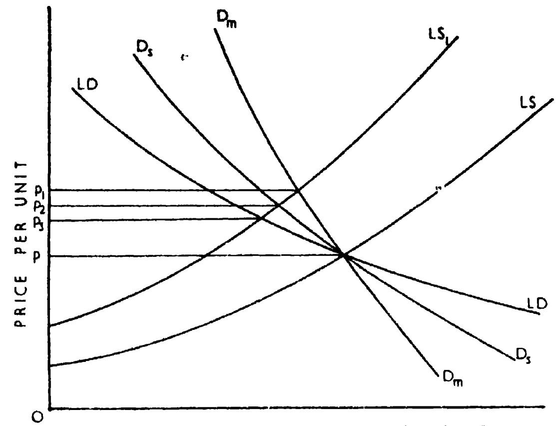


DIAGRAM 51

were described on pages 128 *seq.* above. We might combine Diagrams 48 and 51 to illustrate the effects of long-run demand and supply adjustments on price behaviour over time. If we did so, we would have to be very careful in interpreting our results, for the calendar time required for demand adjustments to take place may differ markedly from that required for supply adjustments.

CONCLUSION

The quantity of a commodity that the firms that make it plan to produce will depend on their production possibilities, the expected prices of the inputs, their objective, and the expected selling price of it. The quantity of this commodity that households will plan to buy will depend on their consumption possibilities (that is, their indifference maps), the expected prices of it and of all the other goods they might buy, their planned consumption expenditures and their objectives. If any one of these should alter, planned purchases and/or planned sales will change also. Economists have concluded that of all these the one that is the least stable over time is the expected price of the product. They have, therefore, isolated the functional relationships between price and planned sales and between price and planned purchases, assuming that all the other things remain unchanged. These relationships are our short-run demand and supply curves respectively. We have shown how price will move to the level at

which these intersect, and how it will behave in response to any given (and unexplained) change in the determinants of demand and supply. In our long-run analysis, we have attempted to explain the nature and direction of the shifts in demand and supply that might occur. The range of choice, for example, from which the firm must choose the production possibilities within which it will confine itself for some time thereafter is not formless: the planning curves illustrate the pattern inherent in it, and from this pattern, *inter alia*, we derived the nature and direction of the changes that would take place in planned output and sales in the long-run. The notion of economies and diseconomies external to the firm described the patterns that input prices might display as firms implemented their long-run decisions.

We shall see later that the concepts of demand and supply provide only rough and simplified approximations to the actual state of economic affairs. These are sufficiently close approximations to reality, however, to enable us to predict the general direction in which relative prices will alter over relatively short periods of time. Even then, they will never tell us that a dozen eggs that is now twice as expensive as a pound of butter will be two-and-one-half times as dear in one month's time; at best, we may be able to say merely that the eggs will be more or less expensive than at present. In this way they are at least as useful to economists as was the scientist's index finger to him before the invention of the thermometer.

The longer the period for which we wish to hazard predictions, the less valuable are these functional relationships between expected price and planned output and purchases, for other things are likely to alter as well, and we can go only a little way towards predicting (or guessing) how these 'other things' will change. It may be that in a dynamic economy, like the United States, Germany, or in the United Kingdom, other relationships than demand and supply — such, for example, as a relationship between expected price and the rate of change of productive techniques — would be more useful. The present state of empirical knowledge, however, makes it at least as hazardous to posit such relationships as to proffer predictions on the basis of demand and supply analysis.

CHAPTER 5

The Purchase Plan of the Firm

In this chapter, we shall describe the purchase plan of the firm and show how it will be revised if there is any change in the data on which it is based. In the next chapter, we shall describe the sales plan of the household and try to discover how it will be revised if the planning data alter. In Chapter 7, we shall show how the relative prices of the things that firms buy and households sell (that is, of productive services) are determined, as firms and households implement their purchase and sales plans respectively. Throughout the last four chapters, we have generally been assuming that the prices of productive services were given and we have concentrated on explaining why the price of a product might change as compared with the prices of other products and as compared with the prices of productive services. In this and in the next chapter, we shall generally assume that product prices remain unchanged, so that our analysis in Chapter 7 will explain changes in the relationship between input prices and between them and product prices. In Chapter 8, we shall combine the analyses of Chapters 4 and 7 to illustrate the simultaneous determination of the prices of products and productive services.

The quantities of each of the inputs that a firm plans to hire or buy and the price per unit that it expects to have to pay for each of them are set out in its purchase plan. In Chapter 2, we described the general form of the purchase plan (*supra*, pages 45–8), and illustrated diagrammatically why a particular purchase plan was chosen by the firm (*supra*, pages 50–70). It will be recalled that we developed the analysis in three steps: first, we delineated the production possibilities open to the going firm (that is, its isoquant map); second, when given these and the prices that it expected to pay for its variable inputs, we showed what quantities of the latter the firm would plan to use to produce each output; third, when, in addition, we were given

the expected selling price of the product, we were able to show what output the firm would plan to produce and sell in each period to maximise its net revenue. In other words, this last step illustrated the simultaneous choice of a purchase plan and of a sales plan (see *supra*, page 69). From Chapter 2, it is clear that the purchase plan that a firm will make will depend first, on the production possibilities that are open to it; second, on the prices of the productive services; third, on the expected selling price of the product; and fourth, on the firm's objective.

We might proceed in this chapter in precisely the same way as we did in Chapter 1. There we showed how the purchase plan of the household would be revised if any one of its determinants altered, all the others remaining the same. We then derived the relationship between the quantity of a good that the household would plan to buy and its price, *ceteris paribus* — that is, the demand for the good — and showed how it would alter if tastes and preferences, or the price of another good, or the planned consumption expenditure should change. It would be both tedious and repetitive to follow the same steps here. Instead, we shall concentrate immediately on deriving the firm's demand for a productive service, and deduce how it will alter when the other data change. We shall suppose first, that the firm is confined to the range of production possibilities that can be achieved by varying only a single input; next, that the quantities of two inputs can be varied, while the quantities of all the others are fixed; and lastly, that all inputs are variable. This progressive widening of the production possibilities that are open to the firm corresponds roughly to a lengthening of the period of operational time: the first two correspond to the short-run and the third and last to the long-run.

THE SHORT-RUN DEMAND CURVE: ONE VARIABLE INPUT

Let us suppose that the quantities of all inputs but *Y* are fixed, so that the firm is confined to the production possibilities lying on *MN* in Diagram 24. The relationship between the input of *Y* and the output of the product that is implicit in this line is shown explicitly in Diagram 25. If we are given the expected selling price of the product, this graph can be translated into a relationship between the input of *Y* and the firm's total revenue. This latter is drawn in Diagram 52: as the input of *Y* is increased,

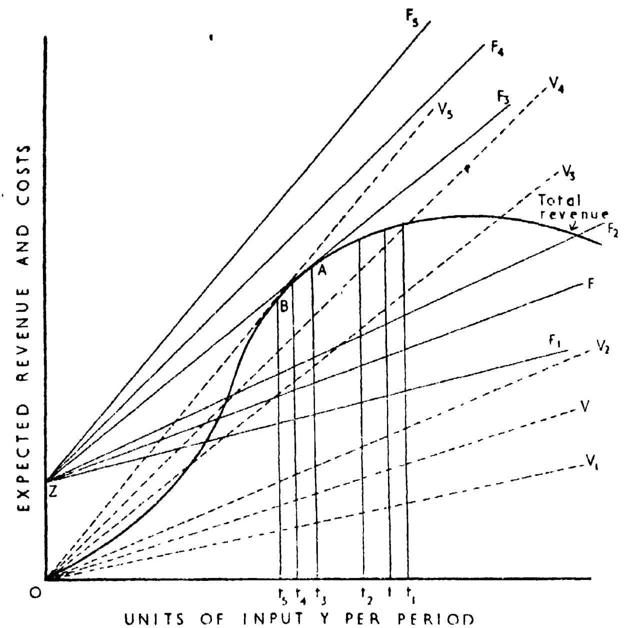


DIAGRAM 52

the firm's revenue for a while rises more than proportionately, and then less than proportionately, because the relationship between the input of *Y* and the output of the product follows the pattern described by the Law of Non-Proportional Returns. If we are now given the price per unit, *y*, that the firm must pay for *Y*, we can graph the relationship between the quantities of *Y* that are used and the total expenditure on their purchase. This is shown by the straight line *OV*, and in our example, where *Y* is the only variable input, it will be the firm's variable cost curve.* Let us suppose that *OZ* represents the expenditure on all the other inputs, the quantities of which cannot be varied: the total costs will then be represented by *ZF*,† which is distant from *OV* at each input of *Y* by the fixed costs *OZ*. The firm, desiring to earn the maximum net revenue per period, will plan to employ *Ot* of *Y*, for only then will the difference between the expected revenue and costs be greatest. The information portrayed in

* Here, the phrase 'variable cost curve' is not used in the same sense as in Chapter 2. In Chapter 2, 'variable costs' meant 'variable costs per unit of output'. Here, it means 'variable costs per period for each level of input'.

† 'Total costs' is not here given the same meaning as in Chapter 2. The difference between the two meanings is the same as that described in the previous footnote.

Diagram 52 is identical with that shown in Diagram 28: the latter shows the behaviour of total revenue and variable and total costs as output is increased, the former the behaviour of these same magnitudes as more of the variable input is used; and in both, the expected selling price of the product, the price of the variable input, and the pattern of production possibilities, are data.

We see from Diagram 52 that the firm in our example, with its given expectations about product and input prices, will plan to buy Ot units of Y per period. If for any reason there is a change in the entrepreneur's expectations about the price he must pay for the variable input Y, then the purchase plan will be revised. If the price falls to y_1, the variable and total cost lines will fall to OV_1 and ZF_1 respectively; since the relation between inputs of Y and output, and the expected selling price of the product, are unchanged, the total revenue curve will remain the same, and the entrepreneur will now plan to buy Ot_1 of Y per period. If the price of Y were to rise from its initial level of y to y_2 per unit, the variable and total cost lines will rise to OV_2 and ZF_2, and the firm will plan to buy Ot_2 of Y per period. By drawing in the appropriate variable and total cost lines, we can discover the purchase plan that the firm would choose at each price at which Y might be bought. It is clear from the diagram that the planned purchases of Y will be the greater the lower is the expected price of Y, and they will be the lower the higher the expected price of Y.

At a price of y_3 per unit, the total cost line ZF_3 just touches the total revenue curve at A, and the maximum net revenue that the firm can earn is zero, and this it can obtain by employing Ot_3 of Y per period. In these circumstances, the sale of the output that Ot_3 of Y produces, when combined with the firm's other factors, will yield a revenue that is just large enough to remunerate all the factors that are being used. If the price of Y were to rise further to y_4, it would still be worth the firm's while to continue in production while the contracts with the other, and fixed, factors remained binding, for by selling the output that Ot_4 of Y produces the firm will earn a revenue that covers its variable costs and makes some contribution to its fixed costs. If the price of Y were to rise to y_5 per unit, the variable cost line would just touch the total revenue curve at B, and the firm will just be indifferent as between continued operation and closure. If it stops

production, it must find its fixed costs of OZ per period; if it hires Ot_5 of *Y* and sells the resulting output, its revenue will fall short of its total costs in each period by the same sum of money.*

This relationship between the planned purchases of *Y* and its expected price that we have deduced from Diagram 52 is called the firm's short-run demand for *Y*. In Diagram 53, we measure the expected price of the variable input *Y* on the vertical axis, and the planned purchases of *Y* on the horizontal axis. When our deductions from Diagram 52 are plotted between these axes, we have the demand curve for *Y*. This curve shows us the quantity of *Y* that the firm would plan to buy at each price at which *Y* might be bought, with its given production possibilities and objective and its contractual obligations to other factors, and its given expectations about the selling price of its product.

The derivation of a firm's demand for a productive service may be illustrated in another way. From the relationship between quantities of input *Y* that are used and total output in Diagram 54 (which is identical with Diagram 25), we can discover the average physical productivity and marginal physical productivity of *Y*.† The average physical productivity (*APP*) is

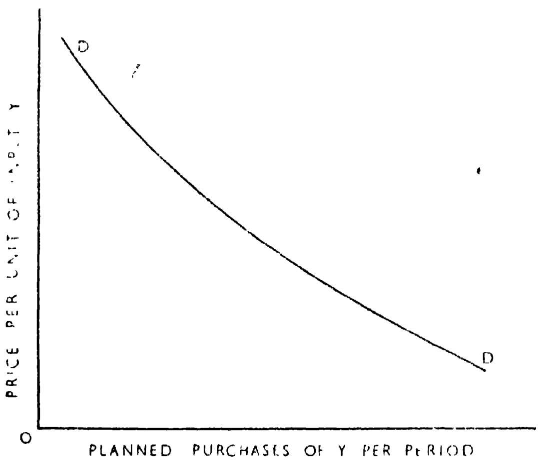


DIAGRAM 53

* Compare pages 70–2 above.

† The label on the vertical axis in Diagram 54 — namely, expected output per period — applies only to the total product curve. For the average product curve, the vertical axis should be labelled: 'output per period per unit of input per period'; and for the marginal product curve, it should be labelled: 'change in output per period per unit change in input per period.'

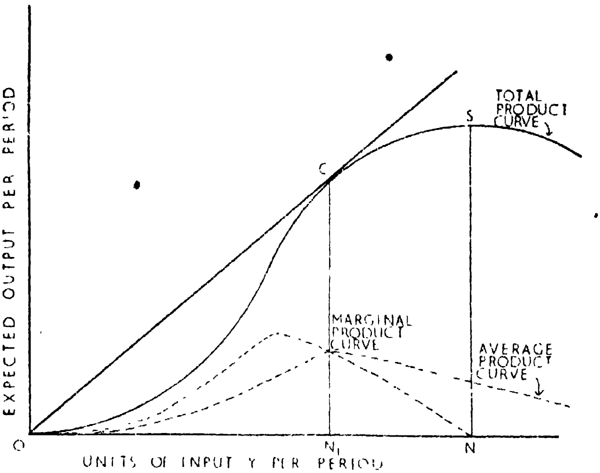


DIAGRAM 54

equal to the total output divided by the number of units of *Y* that are used to produce it: thus, the *APP* of each unit when *ON* of *Y* are being used is equal to *NS*/*ON*, or the slope of the line *OS*. When this calculation is made for each input of *Y* and the results plotted and joined together in Diagram 54, we have the average physical productivity curve of *Y*. The average physical product per unit of *Y* rises for a while as more of *Y* is used, reaches a maximum and then declines. This is a consequence of the shape of the total product curve, which is explained in turn by the Law of Non-Proportional Returns. The marginal physical productivity (*MPP*) of a unit of *Y* is the increase in the total product caused by the use of that unit: if, for example, when 10 units of *Y* are employed the total product is 200, and when 11 units are used the total product is 209, then the marginal physical product of the eleventh unit of *Y* is 9 units of the product. When we calculate the *MPP* at each input of *Y* in Diagram 54 and graph our results, we get the marginal physical productivity curve of *Y*. This curve has the same shape as the *APP* curve and it cuts the latter at its maximum point.*

* Why this is so can be shown geometrically. The *APP* at any input of *Y* is equal to the slope of the line drawn from the origin to the point on the total product curve at that input. In Diagram 54, *APP* is at a maximum when ON_1 of *Y* is being used, for the line *OC* is steeper than any other straight line drawn from the origin to the total product curve at any other input of *Y*. At ON_1, the *MPP* is equal to the slope of the tangent to the total product curve at *C*. Since this tangent must coincide with *OC*, at this input of *Y* the marginal and average physical productivities must be the same.

The average and marginal contributions which different quantities of *Y* make to the firm's revenue can be calculated by multiplying the *APP* and *MPP* of each unit of *Y* by the expected selling price of the product. These average and marginal *revenue* productivity curves are drawn in Diagram 55: they have the same general shape as the average and marginal physical productivity curves in Diagram 54, and for the same reasons. If the firm expects to buy the input *Y* at a price of *y* per unit, it will plan to employ *Ot* of *Y*. This corresponds to the input *Ot* of *Y* in Diagram 52: there, at the input *Ot*, the slope of the total revenue curve (which is equal to marginal revenue productivity of the *Ot*-th unit of *Y*) is the same as the slope of the variable or total cost line (which is equal to the price per unit of *Y*). In Diagram 55, the equality between marginal revenue productivity and the expected price of *Y*, when net revenue is at a maximum, is shown explicitly. We may confirm more directly that net revenue per period is greatest when *Ot* of *Y* is being used: for marginal revenue productivity measures the amount by which total revenue increases when the input of *Y* is increased by one unit; the amount by which the firm expects its total costs to rise if one more unit of *Y* is employed is equal to the price of *Y* per unit. By employing an extra unit of *Y* beyond *Ot*, it is clear from Diagram 55 that the firm would add more to its costs than it would to its revenue. By doing so, the expected net revenue per period would be reduced. Similarly, if the firm used one unit

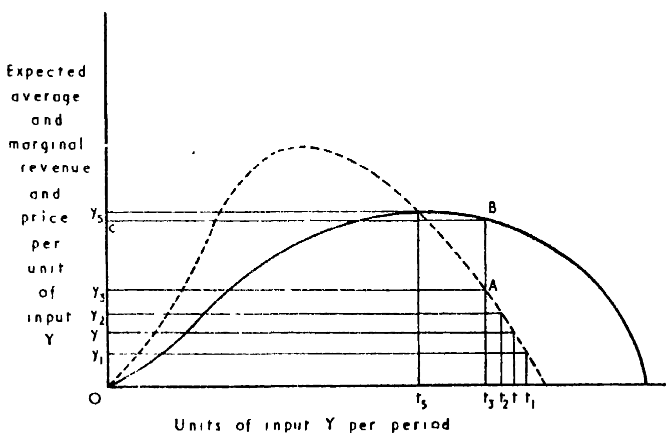


DIAGRAM 55

less of Y than Ot, more would be subtracted from revenue than from costs: by reducing the quantity of Y that is used, therefore, the expected net revenue would be reduced also. If the expected net revenue would be less at inputs of Y greater and less than Ot, then it must be at a maximum when Ot of Y is being set to work with the firm's other factors in each period.

We can show in the same way that at an expected price of y_1, the firm will plan to use Ot_1 of Y; at y_2, Ot_2, and so on. At an expected price of y_3, the firm will employ Ot_3: when it is doing this it will earn a total revenue of Ot_3BC ($= Ot_3$ units of Y multiplied by the average revenue productivity of each unit, t_3B), and its total variable cost will be Ot_3Ay_3; the difference between these, y_3ABC, will just be equal to the firm's fixed costs. If Y is expected to cost y_5 per unit, the firm will be indifferent as between continued production and closure, for the total revenue it would expect to earn when using Ot_5 of Y would be just sufficient to pay for Ot_5 of Y at y_5 per unit. It follows, then, that that part of the marginal revenue productivity curve which lies below the maximum point of the average revenue productivity curve in Diagram 55 is the firm's demand curve for Y, and is in all respects identical with the demand curve drawn in Diagram 53.

It must be emphasised that Diagram 55 is wholly derived from data portrayed in Diagram 52, so that it can add nothing to our knowledge. We may, if we wish, describe the firm's purchase plan for input Y as that which equates the marginal revenue productivity of Y with its price. If we do so, however, we are merely saying in another way that the firm chooses that purchase plan which, *ceteris paribus*, promises the maximum net revenue per period. If we question the empirical validity of the former statement — that is, whether or not a firm consciously strives to equate the marginal revenue productivity and price of the input — we are questioning the validity of the latter.)

We have so far described the direction in which the planned purchases of the single variable input Y will alter if there is a change in its expected price: the higher the price the smaller the purchase, and vice versa. Firms' inputs may be classified according to the extent to which the quantities of them that are bought will vary following changes in their expected prices. This responsiveness of the planned purchases of an input to changes in

its expected price is called the price elasticity of demand for the input. The way in which the planned purchases of Y respond to a change in its price is illustrated by the demand curve in Diagram 55. The strength of the response of the quantity demanded to changes in the price is measured by dividing the proportionate change in quantity demanded by the small proportionate change in price. In Diagram 55, if the price were Oy, Ot would be bought, and if it were Oy_1, Ot_1 would be bought. The price elasticity of demand for Y in this range is, then:

$$E_d = -\frac{tt_1}{Ot} \Big/ \frac{yy_1}{Oy}.$$

Mathematically, the price elasticity of demand for an input will be negative; conventionally, we speak of it as if it were positive. The price elasticity of demand may have any value between zero and infinity. If we wish to speak of this elasticity qualitatively, we shall adopt the same conventions as for the price elasticity of demand for a product (see *supra*, page 36). Its meaning and usage are subject to qualifications similar to those set out on pages 36–7 above in Chapter 1.

The price elasticity of demand for Y in our example, where Y is the only variable input, clearly depends on the shape of the marginal revenue productivity curve, which has the same shape as the marginal physical productivity curve, which in turn depends on the shape of the total product curve. The ultimate explanation of the elasticity of demand for Y must then lie in the pattern of production possibilities open to the firm when Y is the only variable factor. The more rapidly does the rate of rise in output diminish as more of Y is used, the less elastic will be the demand for Y, and vice versa.

The relationship that we have called the demand for input Y shows us the quantity of Y that the firm, in our example, would plan to buy at each price at which Y might be bought, when the production possibilities that are open to it, the contractual obligations that fix the quantities of all the other factors, its objective, and the expected selling price of the product, are all given. We shall now briefly examine what would happen to the demand for Y were any one of these to alter.

First, the effects of a change in the production possibilities. If the quality and kind of the fixed factors had been different

from what they are in our example, we should have a different relationship between total output and inputs of *Y*. If the firm had had more of the same fixed factors, the total, average and marginal physical productivities of each quantity of *Y* would have been greater than in Diagram 54, and the demand curve for *Y* would have been to the right of its position in Diagram 55, and vice versa. If the firm had had other kinds of fixed factors at its disposal, we can say little more than that the total, average and marginal productivities of *Y*, and therefore the demand for *Y*, would have been different from what they are in our diagrams.

Second, the effects on the demand for *Y* of a change in the firm's objective. The market demand for *Y* is what it is in our example because we have assumed, *inter alia*, that the firm strives to maximise its net revenue per period. If the firm sought to cover its total variable costs in each period, its demand curve for *Y* would be the falling portion of its average revenue productivity curve. If its aim were to cover its total costs of production in each period, then the demand curve for *Y* would be a curve lying directly below the falling part of the average revenue productivity curve and asymptotically approaching it as the input of *Y* is increased.*

Third, the effects of a change in the expected selling price of the product. If the price of the product rises, the marginal revenue productivity will be greater than it was before at each input of *Y*, for to obtain it the marginal physical productivity (which is unchanged) is being multiplied by the higher expected selling price of the product. Each point on the demand curve for *Y* in Diagram 55 will move due northwards, so that the new demand curve for *Y* will lie to the right of its original position. Conversely, if the expected selling price of the product falls, the demand for *Y* will fall, and the firm will plan to employ less of *Y* at each price than before.

Lastly, the effects of a revision in the firm's contractual arrangements with its fixed factors. We shall deal later in this chapter with the consequences of a change in the quantity and quality of the fixed factors that widens the range of production possibilities open to the firm. The only contractual revision that

* Compare page 80 above, where the effects on the supply of the product of a change in the firm's objective are examined.

will not alter these latter is one that affects only the size of the fixed costs. If the firm has not been covering its fixed costs for some time, the fixed factors may voluntarily accept a reduction in their rewards to enable it to remain in business. Or the firm may have gone into voluntary liquidation and its fixed factors may have been bought by another firm at their current valuation: the firm that bought them, therefore, will have lower fixed costs per period. Revisions of this kind, however, will have no effect whatsoever on the market demand for the single variable input: for provided the quantity and quality of the fixed factors remain unchanged, the demand for the input is in no way dependent on the fixed costs. A change in fixed costs arising solely from a change in the rewards paid to the fixed factors will, however, alter the length of time for which the firm's market demand for *Y* can be maintained.

THE SHORT-RUN DEMAND CURVE: TWO VARIABLE INPUTS

The production possibilities open to a going firm with two variable inputs are illustrated in Diagram 21. Given the prices of the two variable inputs, *X* and *Y*, we can draw iso-cost lines each showing the different combinations of *X* and *Y* that can be bought with some sum of money. When the iso-cost lines are superimposed on the isoquants, as in Diagram 27, the points of tangency between them show the maximum output that can be obtained for each sum spent on the variable inputs — or, in other words, the minimum variable costs of different outputs. When these 'minimum cost' combinations are joined together, we have the expansion path, and from this we can derive the relationship between output per period and variable costs that is graphed in Diagram 28. Given the expected selling price of the product, and the firm's desire to maximise its net revenue, we can discover the output that the firm will plan to produce and sell in each period. And knowing the output that promises the maximum net revenue, we can discover the quantities of *X* and *Y* that the firm would use to produce it by referring back to Diagram 27. With all this we are familiar from Chapter 2: our problem now is to discover how the quantity of either of the variable inputs that the firm uses to produce its net-revenue-maximising output will vary as its price changes — that is, to derive the demand curves for *X* and *Y*.

In Diagram 56, OE_1 is the expansion path at the initial prices of X and Y. Let us suppose now that the price per unit of input X falls, the price of Y remaining the same. A larger quantity of X can now be bought for each sum of money, so that each iso-cost line in the diagram will be rotated anti-clockwise about the point where it cuts the vertical axis: that is, L_1M_1 swivels to L_1N_1, L_2M_2 to L_2N_2, and to on.* The expansion path, now that X is relatively cheaper, will be OE_2. In Diagram 57, we have drawn the relationships (that are implicit in the expansion paths OE_1 and OE_2) between total variable costs and output. It can be seen that as a result of the fall in the price of X, the total variable cost curve shifts to the right, for each iso-cost line in Diagram 56 now touches a higher isoquant — or, in other words, the variable cost of each output is now less than before. The line OR in Diagram 57 shows the relationship between output and total revenue at the given expected selling price of the product. At the initial prices of the inputs X and Y, the firm will plan to produce and sell OS_1 of its product per period, and from Diagram 56 we can read off the quantities of X and Y that it will plan to employ to produce this output; at the new price for X, when the total variable cost curve is TVC_2, the firm will plan to sell OS_2 per period, and from Diagram 56 we can discover

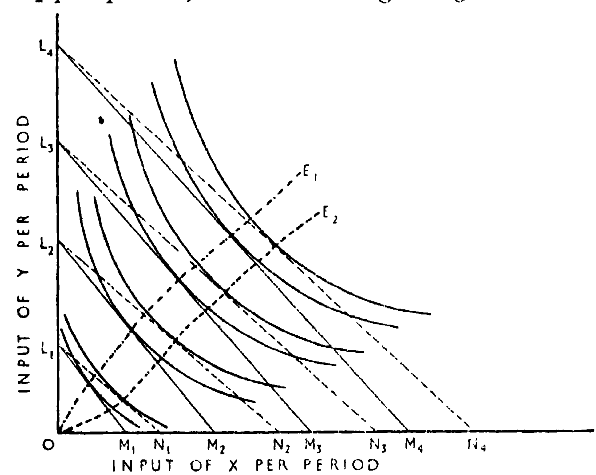


DIAGRAM 56

* To keep the diagram simple, we are assuming that the fall in the price of X is such that each new iso-cost line touches one of the isoquants that are already drawn in the diagram.

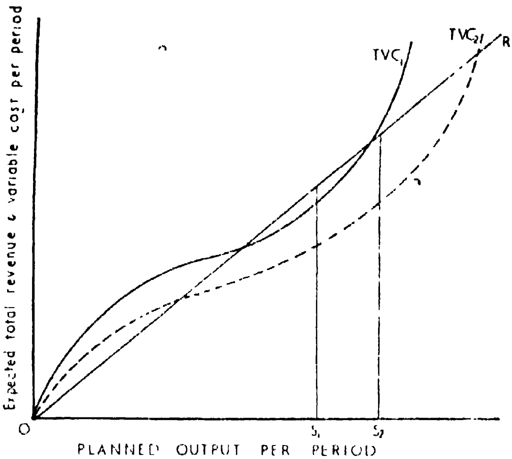


DIAGRAM 57

what quantities of X and Y it would plan to use when doing so. It is clear that the firm will employ more of X at its new and lower price. In a precisely similar way, we can discover what quantity of X the firm would plan to buy at each other price at which X might be bought, and so obtain the demand for X. We will find that the planned purchases of X will vary inversely with its expected price.

The price elasticity of demand for X will be the greater the further is the new expansion path OE_2 to the right of OE_1. The extent of the shift in the expansion path will depend on the shape of the isoquants, and this can be confirmed from the figures in Diagram 58. In Figure (*a*), the marginal rate of technical substitution of X for Y (or of Y for X) falls off rapidly* — that is, the isoquants are highly convex when viewed from the origin; when the price of X falls, the iso-cost line shifts from LM to LN, and the firm will redistribute its expenditure on the two inputs

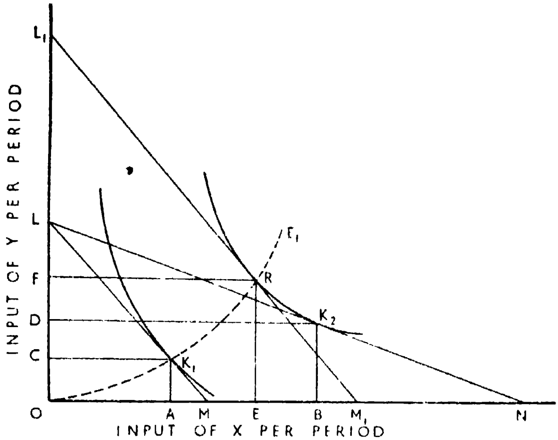


DIAGRAM 58 (*a*)

so as to buy more of both. When a fall in the price of one of the variable inputs causes more of both to be bought, we say that the two inputs are complementary. In Figure (*c*), the isoquants are relatively flat — that is, the marginal rate of technical substitution of *X* for *Y* (or of *Y* for *X*) declines slowly; when *X* becomes relatively cheaper, the firm will re-allocate its expenditure between the two inputs in such a way as to buy more of *X* and less of *Y*. When this occurs, we say that the two inputs are substitutes for one another in production. Figure (*b*) lies between

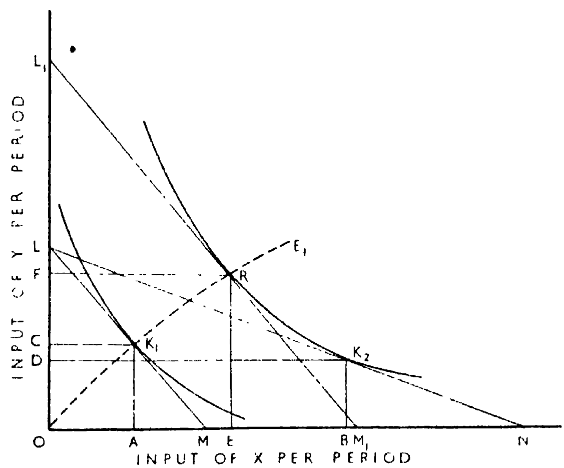


DIAGRAM 58 (*b*)

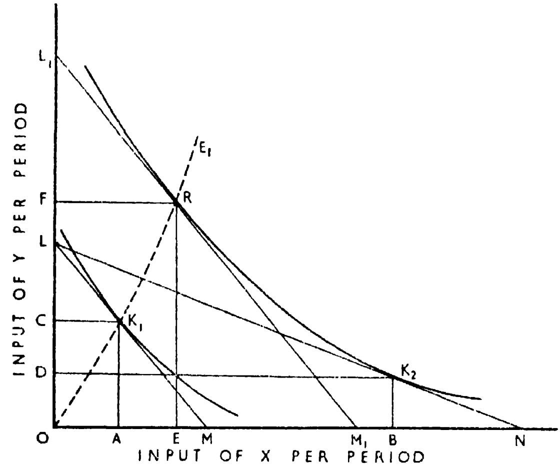


DIAGRAM 58 (*c*)

Figures (*a*) and (*c*); in it, the inputs *X* and *Y* are neither mainly substitutes nor mainly complements.* We conclude that the price elasticity of demand for *X* will be the greater the more easily can *X* be substituted for *Y* in production, and vice versa.

* The maximum output that can be obtained by spending some given sum of money on *X* and *Y* at their initial prices is achieved when *OA* of *X* and *OC* of *Y* are bought. When *X* becomes relatively cheaper, the output-maximising combination of *X* and *Y* that can be bought with the same sum of money is *OB* of *X* and *OD* of *Y*. This change in the output-maximising (or cost-minimising) combination of *X* and *Y* can be regarded as a movement from K_1 on the lower isoquant to K_2 on the higher: and this, in turn, may be viewed as the resultant of two forces. First, now that the price of *X* has fallen, the firm can obtain a higher output for the same expenditure on inputs. It is as if the relative prices of *X* and *Y* had not altered, and planned expenditure had risen. The iso-cost lines L_1M_1 in Diagram 58 illustrate the operation of this force, and its direction and strength are represented by the portion K_1R of the expansion path OE_1. We shall call it the expansion or scale effect: if it operated alone, the firm would plan to buy *CF* more of *Y* and *AE* more of *X*. Second, the firm must next decide how the increase in its purchasing power, now that *X* is relatively cheaper, is to be distributed over the inputs *X* and *Y*. The firm will always buy more of *X* and less of *Y*. This second force is called the substitution effect; it acts along the higher isoquant and its direction and magnitude are represented by RK_2. By itself, the substitution effect would induce the firm to buy *EB* more of *X* and *FD* less of *Y*. The change in the planned purchases of *Y* is then equal to *CF* (the scale effect) minus *FD* (the substitution effect). In Diagram 58(*a*), the former is much larger than the latter; in Diagram 58(*c*) the latter is larger than the former. For input *X*, the scale and substitution effects both work in the same direction. We may say then that the inputs *X* and *Y* are complements when the scale effect of a relative fall in the price of *X* on the planned purchases of *Y* exceeds the substitution effect, and that they are substitutes when the substitution effect on *Y* is greater than the scale effect.

The reader should compare this note with pages 21–3 above. The fact that the isoquants become progressively further apart as we move north-eastwards in the diagram means that the scale effect on *X* will never be negative — that is, that we have no 'Giffen' inputs.

The derivation of the demand curve for input X may be described in another way. In Diagram 59, two isoquants are drawn; the lower shows the combinations of X and Y that would promise 'a' units of the product per period, and the higher those that would yield $(a+1)$ units per period. Let us suppose that the firm is initially employing OR of X and OT of Y to produce 'a' units in each period. The output per period can be increased by one unit by using RS more of X and the same quantity of Y, or by employing TW more of Y and the same quantity of X. We shall call RS the marginal input of X (MI_x), and TW the marginal input of Y (MI_y).* The marginal rate of technical substitution between X and Y in the range CD of the higher isoquant is:

$$MRTS_{x,y} = BC/BD = RS/TW = MI_x/MI_y \qquad (1)$$

We know (see *supra*, Chapter 2, page 63) that when the costs of producing $(a+1)$ units per period are at a minimum, the marginal rate of technical substitution of X for Y is equal to the ratio of the price of Y to the price of X, that is:

$$MRTS_{x,y} = MI_x/MI_y = p_y/p_x \qquad (2)$$

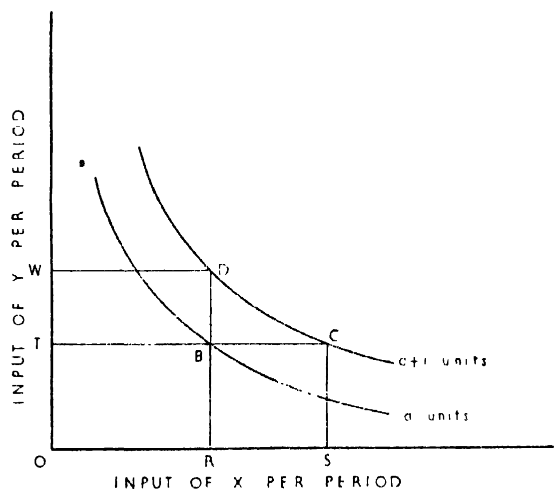


DIAGRAM 59

* The concept of 'marginal input' is borrowed from T. Scitovsky, *Welfare and Competition*, London, George Allen & Unwin, 1952, pages 122-6.

Now, the marginal input of a factor is the reciprocal of its marginal productivity: for the marginal productivity of X, for example, is equal to the increase in the output per period that results from the use of one more unit of X, and the marginal input of X is the number of units of X required to increase the output per period by one unit. Thus, if MI_x is equal to 2 units, then MP_x is equal to $\frac{1}{2}$. We can then translate equation (2) above into a relationship between the prices of inputs and their marginal productivities: when the costs of producing $(a+1)$ units per period are being minimised, then:

$$MI_x/MI_y = MP_y/MP_x = p_y/p_x \tag{3}$$

That is,

$$p_x/MP_x = p_y/MP_y \tag{3a}$$

This equation will be true at each point on the firm's expansion path and at each output on its variable cost curve.

The relationship between marginal inputs and input prices, shown in equation (2) above, can be rewritten as follows:

$$p_x \, . \, MI_x = p_y \, . \, MI_y \tag{4}$$

Each side of this equation is equal to the amount by which the total variable cost would rise if output were increased from 'a' units to $(a+1)$ units per period — that is, each side is equal to the marginal cost of producing the $(a+1)$-th unit. Hence:

$$p_x \, . \, MI_x = p_y \, . \, MI_y = MC \tag{5}$$

We know, however (see *supra*, Chapter 2, pages 74–5), that when the expected net revenue per period is at a maximum, the marginal costs of production will be just equal to the selling price of the product: that is:

$$p_x \, . \, MI_x = p_y \, . \, MI_y = MC = p \tag{6}$$

i.e.

$$p_x/MP_x = p_y/MP_y = p \tag{6a}$$

or

$$p_x = p \, . \, MP_x = \text{marginal revenue productivity of input } X \; (MRP_x) \tag{6b}$$

and

$$p_y = p \, . \, MP_y = MRP_y.$$

When the net revenue per period is a maximum, then, the

quantity of each variable input that the firm employs will be such that its marginal revenue productivity is just equal to its price. If the price of X falls, the firm will adjust its employment of both X and Y until its net revenue is again being maximised and the marginal revenue productivity of each input is again equal to its price. The demand curve for X will be its marginal revenue productivity curve, subject to the same qualifications as were set out on page 145 above. It will slope downwards to the right — that is, the planned purchases of X will vary inversely with the price of X — because of the patterns that are inherent in the production possibilities, and its elasticity will depend on the shape of the isoquants.

The relationship that we have called the demand for input X shows us the quantity of X that the firm would plan to buy at each price at which X might be bought, when the firm's production possibilities, its objective and contracts with the fixed factors, the price of the other variable input(s) and the expected selling price of the product, are all given. If any one of these should alter there will be a change in the demand for X. If the price of Y should fall, and if X and Y are mainly substitutes for one another in production, then the demand for X will fall — that is, the firm will now plan to buy less of X at each price than before. If Y becomes relatively cheaper, and if X and Y are mainly complements, the demand for X will rise, for the firm would now plan to buy more of X at each price. It is unnecessary to describe how a change in any one of the other determinants would affect the demand for X, for a change in the range of production possibilities, or in the firm's objective, or in the price of the product, will shift the short-run demand curve for X when X is only one of the variable inputs in the same direction as it shifts the demand curve for X when it alone is variable (see *supra*, pages 146–8).

THE LONG-RUN DEMAND CURVE

The long-run demand curve may be derived in the same way as the short-run demand curve Given the entrepreneur's knowledge of the techniques of production, and his expectations of the prices he must pay for all the relevant productive services, his planning or long-run average total cost curve can be drawn, as in Diagram 37. Implicit in each point on the planning curve

is a particular combination of productive services — namely, that which promises the lowest average cost per unit for that output. When the expected selling price of the product is known, the planned output of the product and the planned purchases of each service are simultaneously determined. Thus, in Diagram 37, when the price of the product is OP_1 per unit, the firm will plan to produce OM_1 per period with the quantities of inputs that give the point R on the planning curve. If the price of one productive service (X) should fall, then the entrepreneur will have a new planning curve lying south and east of the old one, and the greater is the relative importance of X in each combination of inputs by which the product might be produced, the further south-eastwards will it lie. If the expected selling price of the product remains unchanged, the firm will plan a larger output per period, and will therefore plan to buy more of X and of those inputs that are complementary to X, and less of those inputs for which X can be substituted.

We would expect the long-run demand for X to be relatively more elastic at each price for X than the short-run demand for it, for the same reasons that the short-run demand for X will be more elastic when X is one of two variable inputs than where X alone is variable. In the previous section, we saw that when the price of X fell, X could be substituted for Y: in the long-run, the possibilities of substitution are wider, for X may then be substituted not only for Y but also for the other productive services that were 'fixed' in the short-run.

THE TOTAL DEMAND CURVE FOR A PRODUCTIVE SERVICE

The demand of an individual firm for a productive service shows the quantity of it that the firm would plan to buy in each period at each price at which it might be bought, given the firm's production possibilities, its objective, the price of each other productive service that it uses or might use, and the price at which it expects to be able to sell its product. The role that firms play in determining the relative prices of the productive services they buy is summarised in the total demand curve for each service. If we wish the total demand curve for a productive service to play this role, then we cannot derive it simply by adding together the individual firm demand curves for it. We can show

why this is so by exploring the implications of a total demand curve that is obtained in this way.

In Diagram 60, we suppose that the short-run equilibrium price of the product is OP per unit and the price of the variable input OW per unit, and that the firms are buying WT of the input to help produce PR of the product per period. Let us now suppose that the price of the variable input falls to OW_1: this will cause a rightward shift in the short-run supply curve of the product to S_2S_2. If all firms assume that the price of the product will remain unchanged, they will together plan to buy W_1T_1 of the input; if they simultaneously implement their purchase plans the output per period will be PU, and the price of the product (assuming that it cannot be stored) will fall to OP_1. At

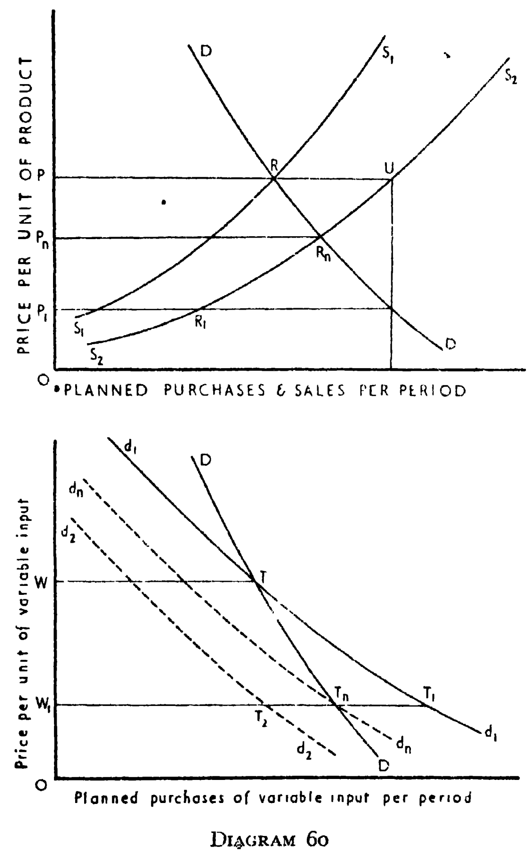


DIAGRAM 60

the price OP_1 for the product and OW_1 for the variable input, firms will plan to produce P_1R_1 of the former with the aid of W_1T_2 of the latter — that is, the whole 'demand curve' for the productive service will shift leftwards to d_2d_2. A cobweb-type cycle will ensue until the product price has reached OP_n per unit and the 'demand' for input d_nd_n; firms will then be producing P_nR_n per period with W_1T_n of the input. If the purchase plans are implemented *seriatim* rather than simultaneously, the supply curve of the product will move gradually to S_2S_2 as the 'demand' for the input falls slowly to d_nd_n, so that the price of the product will fall directly to OP_n and the planned purchases of the input to W_1T_n. In this new position, the purchase and sales plans of the firms will again be consistent with one another. If all points such as T and T_n are joined together, we have the total demand curve for the productive service.

We are already familiar with this kind of problem: in Chapters 2, 3 and 4 we saw that while each firm may make its sales plan on the assumption that the price of each input is given, one consequence of all firms implementing their plans would be to change the relative prices of some or all of their inputs. Here, we see that if all firms lay their purchase plans on the assumption that the price of the product is a datum, one result of all firms putting their purchase plans into effect would be to alter the relative price of the product. In both cases, that which is a datum or constant for the individual firm in an industry is a variable when all firms are taken together. In Chapter 2, we defined the short-run total supply curve in such a way that at each price the sales and purchase plans of the firms in the industry were consistent with one another — that is, in such a way that we could 'move along' it in predicting the probable direction of changes in relative prices. Here, we shall define the total demand curve for an input in a similar way, so that at each price of the input, the purchase and sales plans of the firms in the industry are consistent with one another. This is shown by the curve DD in Diagram 60.

It is clear from Diagram 60 that the short-run total or industry demand curve for a productive service (DD) will be less elastic at each price at which the service might be bought than any of the 'total demand' curves (d_1d_1, etc.) that were obtained simply by adding together the individual firm demand curves. The

elasticity of the total demand curve will depend, then, not only on the degree to which each firm can substitute the input in question for others as it becomes relatively cheaper, but also on the elasticity of demand for the product. If the demand for the product is less elastic than in Diagram 60, then T_n will lie further to the left of the d_1d_1-curve, and the DD-curve will be less elastic also. Conversely, if the demand for the product is more elastic at each price than in the diagram, the total demand curve for the productive service will be more elastic also — that is, T_n will lie nearer to the d_1d_1-curve.

The long-run total demand curve for a productive service may be derived in a manner analogous to that illustrated in Diagram 60. We shall find that the long-run total demand curve will be less elastic at each input price than the 'total demand curve' that is obtained by summing the long-run demand curves for the input of all the firms that might plan to use it over the long period. We shall find also that the price elasticity of the long-run total demand curve will vary directly with the price elasticity of demand for the product.

THE DEMAND FOR A DURABLE GOOD

We have so far confined our attention to the demand for a productive service — that is, for the services rendered by men, land, machines or buildings. Firms demand these services because households (or other firms) demand the products that they help to produce, and in this chapter, we have described the precise manner in which the demand curve for a productive service is derived from the demand curve for the product. Our analysis in this chapter would suffice if firms were always able to buy the services rendered by factors of production in such quantities as they desire. This is true, for example, of labour-service, for any contract by a firm to buy a worker, and so pre-empt his services for all time thereafter, is not legally enforceable. Households frequently own land and sell only its services to firms; and occasionally firms may be able to 'rent' or lease machines, as in the boot and shoe industry. Generally, however, if a firm wants the services of a machine, building, or other durable good, it must buy the good itself — that is, rather than buy the *flow* of services per period it must buy the *reservoir* from which it stems. We shall conclude this chapter, therefore, by

describing the derivation of the firm's demand for a durable good.

In Chapter 3 (*supra*, pages 90–2) we showed (albeit rather superficially) how a firm, when making its long-run plan, might calculate the cost of using the machine in each period of its life. This calculation must now be expressed more precisely. Let us suppose that the machine can now be bought for £P; that its expected life is n periods, and that the firm borrows the money to buy the machine at a rate of interest of i per cent per period. We shall assume also that the costs of operating the machine are zero and that the entrepreneur desires to distribute the cost of the machine equally over all the periods of its life. In each period, therefore, the entrepreneur must set aside a sum of money equal to $(d+P\,.\,i)$, where d is that period's contribution towards recouping the initial price of the machine, and $P\,.\,i$ is the amount of interest that has to be paid in that period on the money he borrowed to buy the machine. Now d, the depreciation per period, will be equal to $P\,.\,i/(1+i)^n-1$,* so that the cost per period of the machine will be $P.i/(1+i)^n-1+P.i$, or $\dfrac{P\,.\,i(1+i)^n}{(1+i)^n-1}$. Given this 'price' per period of the productive service rendered by the machine, and the price of each other productive service, the long-run average cost curve can be drawn, as in Diagram 37. If the expected selling price of the product is OP_1, the firm will plan to produce OM_1 units of output per period over the long-run, with the number of machines and the quantities of other services implicit in the point R on the long-run average cost curve. If the price of the machine should fall, then *ceteris*

* The sum of money that is set aside in the first period of the machine's life (d) may be lent to another entrepreneur and we shall assume that he would pay interest at i per cent per period until the lender requires repayment. The same will be true of the sum, d, set aside in each subsequent period. The value of d must be such, then, that by the end of the period n, the firm will have accumulated a sum of money equal to £P; that is:

$$d(1+i)^{n-1}+d(1+i)^{n-2}+\ldots+d-P,$$

where $d(1+i)^{n-1}$ is the value which the £d lent at the end of period 1 will have reached at the end of period n, and similarly for each other term. This is a geometric progression, and when it is summed we have:

$$\frac{d\{(1+i)^n-1\}}{(1+i)-1}=\frac{d\{(1+i)^n-1\}}{i}=P,$$

$$\text{or } d=\frac{P\,.\,i}{(1+i)^n-1}$$

paribus, the 'price' per period of its services will fall also. As a consequence, the planning curve will move south-eastwards and the entrepreneur will now plan to produce an output larger than OM_1 per period, with more machines and other services. In this way, by drawing the planning curve appropriate to each price of the machine, we can derive the firm's demand curve for the machine, and it is clear that the number of machines demanded will vary inversely with their price. The elasticity of the firm's demand for the machine will depend on the ease with which its services can be substituted for the services rendered by other factors.

The demand curve for a durable good may be derived in another way. The planning curve shows us the minimum unit costs of production of each output: hence, at each output on the curve the ratio between the price of each productive service and its marginal productivity will be the same for all services (see *supra*, pages 154–5). At the output OM_1 per period, which promises the maximum net revenue when the expected selling price of the product is OP_1 per unit, the price of each service will be equal to its marginal revenue productivity (see *supra*, page 154). When the firm is planning to produce OM_1 per period with the number of machines and other services implicit in the point R, then the 'price' of the machine's services per period must be equal to its marginal revenue productivity; that is:

$$\frac{P \cdot i \cdot (1+i)^n}{(1+i)^n - 1} = MRP \qquad (1)$$

$$\text{or,} \quad P = \frac{MRP}{i}\left(\frac{(1+i)^n - 1}{(1+i)^n}\right) \qquad (2)$$

The right-hand side of this equation, however, is merely the present value of the marginal revenue productivity of machines (when the planned output is OM_1 per period) over each period of their lives, for the present value (V) of sums of money of MRP accruing in each of the n periods for which the machine will last is:

$$\frac{MRP}{(1+i)} + \frac{MRP}{(1+i)^2} + \frac{MRP}{(1+i)^3} + \dots + \frac{MRP}{(1+i)^n}.$$

The sum of this geometric progression is:

$$\frac{MRP}{1+i}\left\{1-\frac{1}{(1+i)^n}\right\}\Big/1-\frac{1}{1+i}=\frac{MRP}{i}\left(\frac{(1+i)^n-1}{(1+i)^n}\right).$$

Equation (2) tells us that when the expected net revenue per period is at a maximum, the present value of the marginal revenue productivities per period of machines will be equal to the price at which they can now be bought. If the price per machine should fall, then the firm would plan to buy as many more machines as were necessary to reduce the marginal revenue productivity per period of machines to the level at which their present value was equal to the new price, for only then would the expected net revenue per period be at a maximum in the new conditions. The elasticity of the demand for machines will depend on the rate at which their marginal revenue productivity per period declines as more of them are used, and this in turn depends on the degree to which the services of machines can be substituted for other productive services as machines become cheaper.

The choice of the number of machines (of some given kind) that the firm will plan to buy is generally made when the entrepreneur is making his long-run plan. The firm's demand curve for a durable good, therefore, relates to the long-run, and it shows us the number of machines that it will plan to buy at each price at which they might be bought, given the entrepreneur's knowledge of the techniques of production, his objective, the price of each other durable good or productive service, and the rate of interest, the expected life of the machine in question and the expected selling price of the product. If the rate of interest that the firm uses in making its calculations should fall, then the present value of the stream of marginal revenue productivities will rise, so that the firm will plan to buy more machines at each price: that is, the whole demand curve for the machine will move to the right. If the rate of interest rises, the demand curve will move to the left. If the firm believes that the rate of technical development, and therefore of obsolescence, will be slower, and so comes to expect the life of each machine to be longer, then the present value of the machine will rise and its demand curve will shift rightwards. Conversely, if the expected life of the machine is shortened, the demand for it will fall. Changes in any

one of the other determinants of the demand for a durable good will affect the demand for it in the same way as they would affect the demand for a productive service.

The total demand curve for a durable good may be derived in a manner analogous to that illustrated in Diagram 60. The total demand will be less elastic at each price than the 'total demand' that would be obtained simply by adding together the demand curves of all the firms that might plan to use that good in the long-run. We shall find also that, *ceteris paribus*, the elasticity of the total demand for a durable good will vary directly with the price elasticity of demand for the product that it helps to produce.

In this chapter, we have described how the total demand for a durable good or productive service is derived from the demands of the individual firms for it. The total demand curve for a factor or factor-service summarises the part that firms play in determining the relationship between the prices of the things that they buy. In the next chapter, we shall study the sales plans of households and try to discover how households, in implementing their sales plans, help to determine the relative prices of the productive services they sell.

CHAPTER 6

The Sales Plan of the Household

In this chapter, we shall describe the sales plan of the household and show how it will be revised if any of the data on which it is based should alter. Households own factors of production and sell their services to firms. The members of a household own their labour, and the price per unit at which they sell their labour-service is called the wage-rate. A household may own durable goods, like buildings or plots of land: the price at which the services that these render are sold is called rent. Finally, a household may have money — that is, the power to acquire goods and services — which it is accumulating, or has accumulated, from its income, or which it has acquired by gift or bequest; the price that firms pay for the use of money per period is called the rate of interest. Here, we shall first study the sale of labour-service by households, then the sale of the services of durable goods, and lastly, the sale of the services of money.

The quantities of each productive service that a household plans to sell and the price per unit that it expects to receive for each of them are set out in its sales plan. If we use 'p' to represent the expected selling price, 'q' to represent the quantity that the household plans to sell, and the suffixes 1, 2, 3 ... n, to designate particular productive services, then the sales plan of the household will have the following general form:

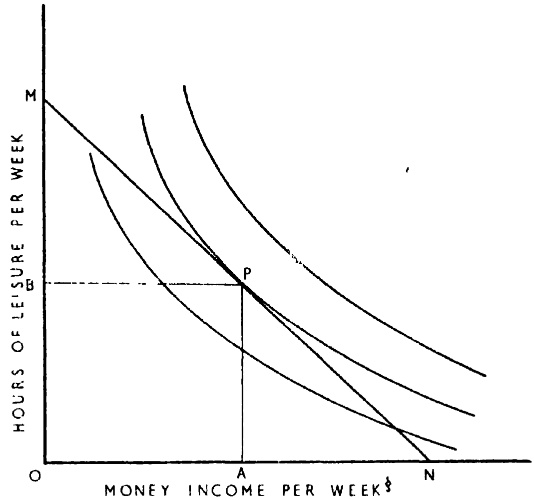
$$p_1 \cdot q_1 + p_2 \cdot q_2 + \ldots + p_n \cdot q_n = y,$$

where 'y' represents the income that the household hopes to receive by implementing the sales plan as a whole. We shall assume that the price at which each productive service might be sold is a datum for the individual household. The problem that faces the household is to decide what quantity of each productive service to sell.

THE SALES PLAN FOR LABOUR-SERVICE

We shall suppose initially that the household owns only one factor of production — namely, one carpenter — and that the sales plan relates to the week that lies ahead. Our problem is to illustrate the choice of sales plan — that is, the number of hours of his labour-service that the carpenter plans to sell — so that (*a*) both the range of choice and the considerations that influence it are made clear, and (*b*) we can deduce how the plan will be revised.

In selling his labour-service, the worker is buying a money-income that can be used to sustain and amuse himself and his family, both now and in the future. The sustenance and amusement, that any given money-income is expected to provide, depends (as we saw in Chapter 1) on the household's tastes and preferences for the goods and services of everyday consumption and the prices that it expects to have to pay for them. A carpenter, however, is not merely a wood-working machine: when not working as a carpenter he may play with his children, dig his garden or watch television. Since there is a limit to the number of hours at his disposal each week, the more hours he sells for income by working as a carpenter, the fewer hours will he have in which to indulge his other interests. How he will dispose of his time between working as a carpenter for money-income, and pursuing these other interests for his own pleasure, will depend on his tastes for each and on his preferences for different combinations of them. These are illustrated in Diagram 61. On the vertical axis we measure the number of hours that the worker will plan to devote to activities other than working for money as a carpenter; for the sake of brevity, we shall follow the common usage and call these hours of leisure, though the 'leisure-time' activities may be both physically and mentally more exhausting than carpentry. On the horizontal axis, we measure money-income: since we are assuming that the tastes and preferences for consumption goods and their prices are given, each sum of money-income will denote a particular bundle of real goods and services. Each point that lies between these axes will represent a particular combination of income and leisure, possessing a certain degree of attractiveness to the members of the household. The relative attractiveness of different combinations of income



§Since all product prices are assumed to be constant, each money income will represent a particular level of real income i.e., the adjectives money and real to qualify income on the horizontal axis are interchangeable.

DIAGRAM 61

and leisure can be seen immediately if we draw indifference curves, each passing through all combinations that the household finds equally attractive. Here, as in Chapter 1 (see *supra*, pages 9–12) both logic and experience help us to determine the shape of these indifference curves. Since of any two combinations of income and leisure, each containing the same quantity of leisure (or of income), the household will prefer that with the larger quantity of income (or of leisure), no indifference curve can run north-eastwards in the diagram. The logically permissible shapes are as shown in Diagram 3, and of these only indifference curves that are convex when viewed from the origin are consistent with the results of introspection and with our knowledge that when the hourly wage-rate rises, the number of hours that an individual worker will work may rise, fall, or remain the same. When all the indifference curves are drawn, we have the household's indifference map, and this illustrates

the household's tastes or desires for money-income and leisure, and its preferences as between different combinations of them.

We shall assume that the objective of the household as a seller of productive services is to maximise its satisfaction. In this pursuit, it is hampered by the limited time at its disposal, and by the wage per hour on which depends the rate at which time can buy money-income. The nature of these limitations can be illustrated in Diagram 61. We shall suppose that the worker must sleep for 8 hours per day, and that the remaining 16 hours per day, or 112 hours per week ($= OM$) are available for work and leisure. If the worker spends all the time at his disposal on 'buying' leisure, he can have OM hours per week, for the price of one hour of leisure is always one hour of time. If the hourly wage-rate is 2s. 6d., and if the worker spends all these hours on buying money-income, he can obtain an income of ON per week. All the combinations of money-income and leisure that the worker might have at this hourly wage-rate will then lie on the straight line MN; of all these, he will choose that which he prefers, and this is shown on the diagram by the point P where MN touches an indifference curve. At P, the household is planning to enjoy OB hours of leisure per week and to use the remaining BM hours to buy a weekly income of OA. The sales plan of the household is then:

$$BM \text{ hours} \times \text{hourly wage-rate} = OA.$$

We have now illustrated diagrammatically the sales plan of a household that is selling only one kind of labour-service. Though in doing so we have added nothing to our knowledge, we have at least clarified the circumstances in which the sales plan will be revised. It should be clear from the preceding pages and from Diagram 61 that the sales plan will be revised if there is any change in the worker's tastes and preferences for income and leisure, or in the prices of the products that the household buys, or in the hourly wage-rate.

If the prices of the goods and services that the household plans to buy should rise, then each sum of money-income will buy less of these things than before. From the household's point of view, then, it is as if the prices of the things it buys remained unchanged while the hourly wage-rate fell. The consequences of

changes in product prices, therefore, will be similar to changes in the opposite direction in the wage-rate.

The possible reactions of a household to changes in the hourly wage-rate are illustrated by the figures in Diagram 62. If the wage-rate were *w*s. per hour, the combinations of leisure and money-income that the household could attain would lie on *MN* and of these the household would choose *P*. If the wage-rate were to rise to w_1s. per hour, the household would choose the sales plan implicit in P_1; if it were still higher at w_2s. per hour, it would choose P_2, and so on. The line joining $MP_1P_2P_3$ in the diagram is analogous to the price-consumption curve, and it shows us how the household would allocate its time between leisure and work at each hourly wage-rate.

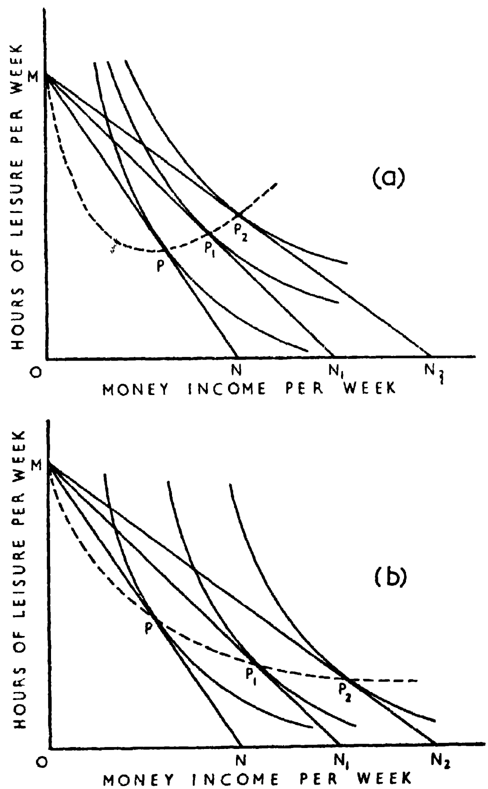


DIAGRAM 62

The relationships between the hourly wage-rate and the number of hours that the worker is willing to work which are implicit in the $MP_1P_2P_3$ ... lines in Diagrams 62(a) and 62(b) are shown explicitly in Diagrams 63(a) and 63(b) respectively. In 62(a) the tastes and preferences of the worker for leisure and income are such that as the wage-rate rises above ws. per hour, the number of hours for which he is willing to work will fall — that is, his supply curve of work will be 'backward sloping' as in Diagram 63(a). If his tastes and preferences are as illustrated in Diagram 62(b) then the number of hours for which he will plan to work will vary directly with the hourly wage-rate; that is, his supply curve of work will be as shown in Diagram 63(b).

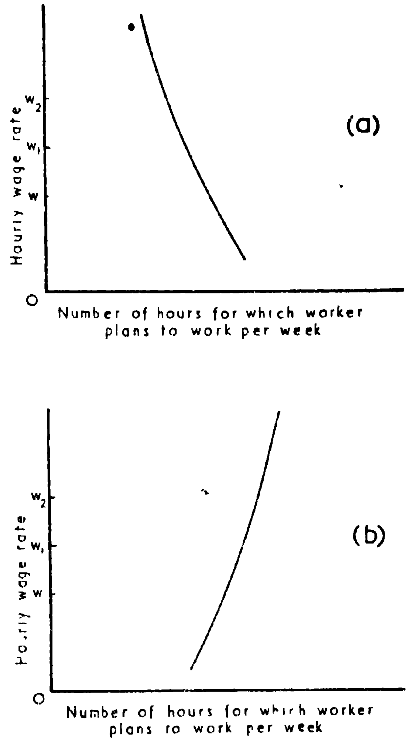


DIAGRAM 63

Thus far we have merely shown that the direction in which the number of hours that the worker is willing to work will alter when the wage-rate changes depends on his indifference map. The characteristics of the indifference map that determine the kind of reaction can be described in another way. In both the figures in Diagram 62, the worker is 'better-off' the higher is the hourly wage-rate, for the points P, P_1, P_2, etc., lie on progressively higher indifference curves. We may think of the rejection of P and the adoption of P_1, when the wage-rate rises from ws. to w_1s. per hour, as a 'movement' along the curve $MP_1P_2P_3$ ... . The 'force' (namely, the rise in the wage-rate) that pushed the worker in this direction can be thought of as being the resultant of two other forces. First, when the wage-rate rises, the worker

may have more income and the same leisure, or more leisure and no less income — that is, he is potentially better-off. It is as if the wage-rate had remained unchanged and the worker had been given a sum of money equal to NQ in Diagram 64. We may, therefore, think of this first force as operating along the line PR: we shall call it the income-effect, because the increase in the level of satisfaction of the household that follows a rise in the wage-rate is as if it had received an income of NQ per week from some source other than its labour. If the income-effect alone operated, the worker would plan to work for BD fewer hours per week. The second component of the movement from P to P_1 is the substitution effect: when the hourly wage-rate rises, each hour devoted to work yields more income — that is, leisure becomes relatively more 'expensive'. The magnitude and direction of the substitution effect will always be positive, for the worker will always tend to substitute income, which his time can now buy more cheaply, for leisure, the price of which has not altered. By itself, the substitution effect would induce the worker to offer DF more hours of his labour-service for sale each week. In Diagram 64(*a*), wherein the indifference map of Diagram 62(*a*) is reproduced, the negative income effect outweighs the positive substitution effect, so that fewer hours are worked when the wage-rate rises. In Diagram 64(*b*), the substitution effect is stronger than the income effect, so that the individual worker's supply curve rises monotonically.

The circumstances in which the number of hours of his labour that the worker is willing to sell in each week will rise or fall when the hourly wage-rate rises can be described in yet another way. Columns (1) and (2) in Table 2 show the relationship between the wage-rate and the number of hours worked: as the wage-rate rises from 30 pence to 33 pence per hour, the worker offers more hours of his labour for sale, as in Diagram 63(*b*); at wage-rates above 33 pence per hour, he will plan to work fewer hours per week, as in Diagram 63(*a*). Now in selling his work (or effort) the worker is buying income. The effort-price that he must pay for one pound of income is the number of hours for which he must work in order to earn it at the ruling wage-rate: thus, when the wage-rate is 30 pence per hour, he must work for 8 hours to earn £1. The amount of income which he demands is equal to the hourly wage-rate multiplied by the num-

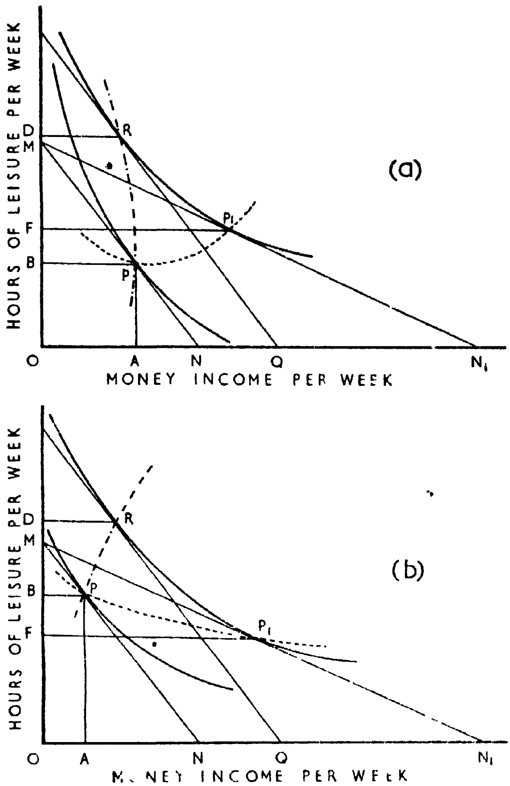


DIAGRAM 64

ber of hours for which he would plan to work: thus, when the wage-rate is 30 pence per hour, he is willing to work for 36 hours — that is, he is demanding an income of 30 × 36, or 1,080 pence per week. The effort-price of a unit of income is calculated for each hourly wage-rate in column (4), and the total income that the worker demands at each wage-rate is set out in column (3) of Table 2. In Diagram 65 we plot the income that would be demanded at each effort-price of income; when the points are joined together, we have the individual's demand curve for income in terms of effort.

TABLE 2

Wage-rate (pence per hour) (1)	*Hours worked per week* (2)	*Total income demanded per week (pence)* (3)	*Effort-price per unit of income (that is, hours per £1)* (4)
30	36	1080	8·0
31	40	1240	7·7
32	44	1408	7·5
33	48	1584	7·3
34	47	1598	7·1
35	46	1610	6·85
36	45	1620	6·67
37	44	1628	6·5

The 'effort-price' elasticity of this demand curve may be measured by the total expenditure method described on pages 37–9 above. Thus, when the effort-price is 8 hours per £1 (that is, when the hourly wage-rate is 30 pence), a weekly income of 1,080 pence is demanded, and the worker's total expenditure of effort in buying this income is 36 hours; when the effort-price falls to 7·7 hours per £1 (that is, when the hourly wage-rate is

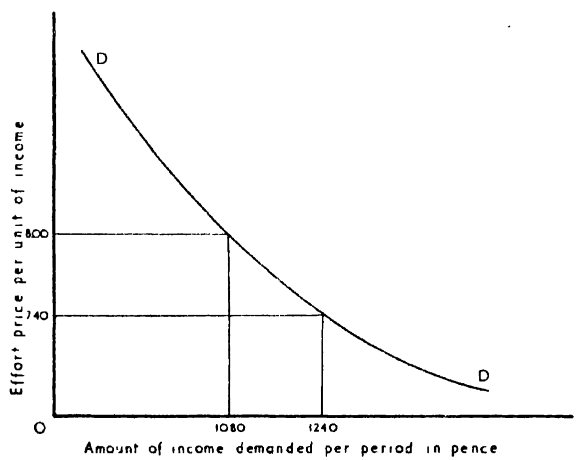


DIAGRAM 65

31 pence), a weekly income of 1,240 pence is demanded, and the worker's total effort expenditure is 40 hours. In this range of the demand curve, the total effort expenditure rises as the effort-price falls — that is, the worker's demand for income in terms of effort is relatively elastic. It is clear that the effort expenditure is merely the number of hours that the worker is willing to work at a particular wage-rate — and these are given in column (2). By comparing columns (4) and (2) of Table 2, then, we can see immediately whether the worker's demand for income in terms of effort is relatively elastic or relatively inelastic: as the effort-price falls from 8 to 7·3 hours per £1 (that is, as the wage-rate rises from 30 to 33 pence per hour), the demand for income in terms of effort is relatively elastic; at effort-prices lower than 7·3 hours per £1 (that is, at wage-rates higher than 33 pence per hour), the demand for income in terms of effort is relatively inelastic.

We may sum up thus far as follows: It has been observed that when the hourly wage-rate rises, some workers offer more hours of their labour for sale and some offer less. From these facts we inferred the shape of the leisure-income indifference curves. The fruit of our indifference analysis was not an explanation of why a worker reacts to higher hourly wage-rates in the way that he does, but rather a classification (under the headings: tastes and preferences for income and leisure, prices of products, etc.) of the different influences that affect his decision. The segregation of the substitution- and income-effects of a rise in the hourly wage-rate, and the concept of the elasticity of the demand for income in terms of effort, merely offer alternative ways in which these same facts can be communicated to others: if, when the wage-rate for carpenters rose from 2s. 6d. to 2s. 9d. per hour, carpenter Bill Smith reduced the number of hours for which he was willing to work per week, we may describe Bill Smith's behaviour in precisely these words, or we may say that for him the income-effect of the increase in the wage-rate outweighed the substitution-effect, or we may say that his elasticity of demand for income in terms of effort was less than unity. All this, however, does not help us to predict how Bill Smith would react if another wage change should occur, for we cannot establish by empirical investigation the precise characteristics of his indifference map, or the relations between his income and substitu-

tion effects, or the elasticity of his effort-demand for income. If we wish to be able to predict the probable consequences of a rise in the wage-rate, we must discover some criteria by which we can recognise whether a worker falls into the group that will work fewer hours, or into the group that will work more hours, per week. To this task we now briefly repair.

Where the wage-rate has been at the level of *w*s. per hour (in Diagram 64) for some time, so that the worker has become accustomed (or reconciled) to the standard of living that it can command, we commonly find that the number of hours worked falls as the wage-rate rises. This tendency will be the stronger the more exhausting is the work that he is doing, and the more numerous are the opportunities for passing leisure-time inexpensively. Where the prevailing wage-rate does not enable the worker to achieve the standard of living to which he aspires, or to maintain the standard to which he has become accustomed in the past, the number of hours that he is willing to work will usually vary directly with the wage-rate per hour. The prospect (or fact) of marriage and children, for example, and the expenses that attend them, may induce this kind of behaviour. Indeed, it is conceivable that more hours may be worked, not merely because the worker aspires to a higher standard of living, but because he aspires to more expensive hobbies.

We have now derived the supply curve of labour-service of the individual carpenter. The total or market supply curve of carpenters' services is obtained by summing together these curves — that is, by adding the number of hours that each carpenter is willing to work at each hourly wage-rate. The shape of the total supply curve of labour-service from those possessing a particular skill will depend on the degree of their preferences for leisure as opposed to income. If the preferences for leisure are strong, we would expect the total supply curve to have the same shape as that drawn in Diagram 63(*a*); if the individual indifference maps are on balance like that illustrated in Diagram 62(*b*), we would expect the total supply curve to be like that drawn in Diagram 63(*b*).

In this chapter so far we have assumed that the worker's current behaviour is circumscribed by past decisions. Some time in the past, he acquired the skill of carpentry. While this decision binds him, he is limited in revising his sales plan to the various

ways in which the total time at his disposal can be allocated between leisure and working as a carpenter; and we have illustrated his choice of that allocation of his time that promises him the maximum satisfaction, given his expectations about the hourly wage-rate and the prices of the products he may want to buy. We shall call an analysis that is so confined a *short-run* analysis, and contrast it with a *long-run* analysis which explores the choice of a sales plan when the worker may choose what skill to acquire. Here, these terms have meanings analogous to those we gave them in Chapters 2 and 3 when we described the sales plans of firms and their revisions. There, in the short-run, the number of firms that might produce a product was given, and no firm could vary the size of its plant Here, in the short-run, the number of workers with a particular skill is given and no worker can change the skill he possesses. There, in the long-run, the number of firms that could produce a particular commodity might vary through entry and exit, and existing firms might change the size of their plants. Here, in the long-run, the number of carpenters might vary through recruitment, resignation and death, and men who are carpenters might become bus-drivers or bricklayers. We shall now describe the choice of a long-run sales plan by households, and show how the long-run supply curve of a particular kind of labour-service may be derived from the manner in which households revise their sales plans as relative wage rates alter.

When a worker makes a long-run sales plan, he is deciding what kind of labour-service to sell — that is, what skill to acquire. Our problem is to describe the range of choice that faces the worker and the considerations that influence his final decision. We shall seek its solution by taking the simplest example, namely, that of a youth who has reached the legal minimum age at which he may undertake full-time employment. The range from which he must choose is merely a catalogue of all the skills or 'occupations' of which he is aware. Some of these may be eliminated by his assessment of his own abilities and potentialities. From the occupations that he feels competent to enter, he will make his choice in the light of (*a*) his and his fellows attitudes towards the type of work — whether manual or mental, monotonous or exciting, hazardous or safe — that each occupation requires, and the conditions in which it must be per-

formed — whether outdoor or indoor, sitting or standing, in lounge-suit or overalls; (*b*) his estimate of the time and cost of preparing himself for each occupation; and (*c*) his expectations of the income per period that he might earn were he to enter each occupation. For many workers, the second of these may drastically narrow the range of choice, for they may neither possess nor be able to acquire the money that is needed to meet the costs of being trained for certain occupations: for them, it is as if the costs of becoming a doctor or lawyer or school-teacher were infinitely large. As economists, we can say little more than that the worker will choose that occupation that he prefers, and we presume that in making his choice he in some way adds together the relative income he would hope to earn were he to pursue each occupation, the relative attractiveness to him of the kind of work it requires, and the relative social esteem in which it is held. A worker who already possesses a skill will be influenced by similar considerations in deciding whether or not to acquire a new skill.

The long-run decision of a worker will be revised if there is any significant change in his estimate of his own capacities, in his attitudes towards the type and condition of work in different occupations, in the relative costs of preparing himself for different occupations, or in the relative money-incomes he expects to earn were he to enter them. We cannot derive a long-run supply curve for the individual worker from the way in which his sales plan would be revised when the relative prices of different kinds of labour-service alter, for in the long-run the worker is choosing between different full-time occupations. Given all the other influences that affect his choice, at one set of relative prices he might decide to become a school-teacher; at another, a carpenter, and at yet another he might plan to become an agricultural labourer. From the manner in which each individual worker will revise his long-run sales plan as relative wage-rates (or relative salaries) alter we can, however, derive the long-run supply curve of carpenters, or of school-teachers, or of workers to any other occupation. We can do this in the following way.

Let us suppose that all the influences that we have listed remain the same, but that the hourly wage-rate of carpenters rises. As a consequence, we would expect more new entrants to the labour market to plan to become carpenters, and some of those

who had previously chosen other occupations to revise their decisions and train as carpenters. Initially, the 'new' carpenters would probably be drawn from occupations that required similar abilities and offered similar conditions to the trade of carpentry. As the wage-rate that carpenters might earn rose further, however, the 'new' carpenters might be drawn from semi-professional or professional occupations — for there is some wage-rate that might induce even professors and surgeons to ply this trade. The higher is the wage-rate that firms are willing to pay for carpenters' services, therefore, the larger the number of workers who will plan to become carpenters, and vice versa, so that the long-run supply curve of carpenters will slope north-eastwards as in Diagram 66.

The elasticity of the supply of carpenters is the responsiveness of the number of carpenters to changes in the relative wage-rate, and it is measured by dividing the proportionate change in the number of workers who are planning to work as carpenters by the small proportionate change in the expected wage-rate. In the absence of special measures by trade unions or professional associations to exclude new entrants, we would expect the long-run supply curve of labour-service to any particular occupation to be relatively elastic, for as the wage-rate that can be earned in it rises, workers will be drawn from other occupations that are held in comparable social esteem and that require similar

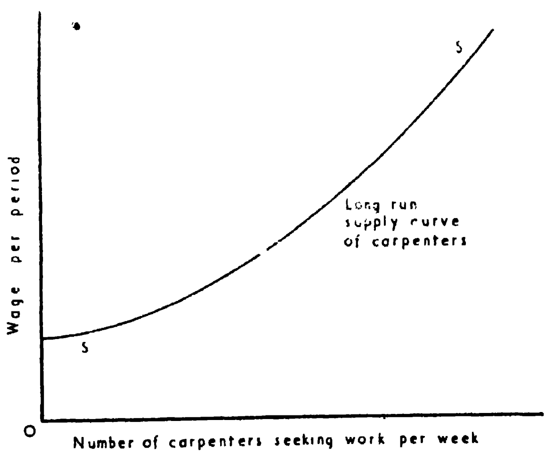


Diagram 66

abilities and training expenses. The elasticity of the long-run supply of labour-service to an occupation will generally be positive, and it will always be greater at each wage-rate than that of the short-run supply curve, for in the long-run the entry or exit of new workers will generally outweigh variations in the number of hours for which each man that is in the trade is willing to work. It is possible that the long-run supply curve of labour to an occupation that lies at the bottom of the 'occupational ladder' may have a negative elasticity over a part of its range: for this to happen, it must be so unremunerative or suffer from such social opprobrium that workers are not drawn into it from occupations that lie immediately above it in the occupational scale even when its wage-rate rises over this range.

THE SALES PLAN FOR THE SERVICES OF LAND

We shall suppose that the household owns a given area of land. The land may yield its services in the production of many different commodities and services: thus, it might assist in the production of corn or wheat or potatoes or of any one of many other agricultural products, or it might be used as a building-site and so assist in the production of housing-services, or it might be used as a pleasure-ground and so help to provide some kind of 'social utilities'. The choice of a sales plan by a land-owner is the choice of the way in which his land is to be used. We shall define the short-run as the period within which land must continue in its present use:* the calendar time for which the short-run will endure will vary widely from one use to another. It may be only one season for land that is being used to grow potatoes or corn, or it may be a century for land on which houses have been built. In the long-run, the kind of production to which the

* The short-run might here have the same definition as we gave it when dealing with labour-service in the previous section. There, the worker had two uses for his time in the short-run: he might use it to buy income by selling his existing skill, or he might use it as leisure. Similarly, the landowner in the short-run might use his land to buy income in its present use, or he might put it to some use — such as a flower garden — that yields some kind of personal satisfaction but no income. It is unlikely, however, that such a definition would ever be realistic, and if it ever seemed to be realistic, the explanation would most likely lie in a 'backward sloping' supply curve for some kind of labour-service. Thus, as the actual (or imputed) price per acre of land that is used to grow wheat rises, a farmer may devote more land to growing flowers for his own pleasure: the reduction in the wheat acreage, however, is a consequence of the farmer's re-allocation of his time between income and leisure.

land contributes may be altered: thus, land that is now growing wheat may have houses built upon it, or buildings that are now on the land may be demolished and the site used as a car park. Our problem, as before, is to derive the short-run and the long-run supply curves of land to each avenue of production, explain their elasticities and list the circumstances in which they will shift to new positions.

In the short-run, as we have defined it, the household cannot choose by whom or for what purpose its land is to be used. Its present use is the result of a past decision which is unalterable in the short-run. The household may have leased its land on an eleven-months' contract to a farmer who is using it to grow corn; or the household may itself be a farm that is 'leasing' its land to itself and using it to grow corn. Until the contract expires or the corn is reaped, the land cannot be diverted from its existing use in response to any change in the rent per acre that land* can earn in corn-production. In the short-run, therefore, the individual household's supply curve of land to each use will be perfectly inelastic; and the total or market supply curve, which is obtained by adding together the individual supply curves, will be perfectly inelastic also.

In making a long-run sales plan, the owner of land is deciding upon the use to which his land shall be put — that is, upon the product genus to whose production it will contribute. The range from which he must choose is merely the list of all the uses of which he is aware and for which his land is suited by its inherent and acquired qualities and its situation. Given the rent (that is, the price) per acre that he expects his land to earn in each use, he will choose that which promises him the greatest reward. He will revise his sales plan if there is any significant change in his expectations of what his land could earn in different uses. We cannot derive a long-run supply curve of land for the individual landowner from these revisions for in the long-run he is choosing between different full-time uses of his land. We can, however, derive the long-run supply curve of land to any particular use from the way in which each individual landowner revises his long-run sales plan as the relationship between the prices that

* In this section, for brevity's sake, we shall use 'land' to mean the services rendered by land. When 'land' is being used to mean the source of these services and where the context does not make this clear, we shall say so.

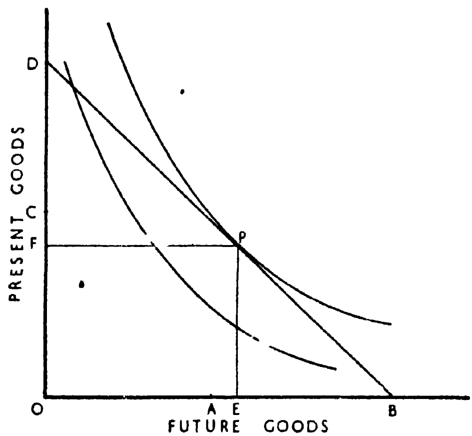
are being paid for land in different uses changes. As the price per acre for which the services of land can be sold to corn production rises, the price per acre in each other use remaining the same, we would expect more and more landowners to plan to devote their land to that use, and vice versa. We would expect the long-run supply curve of land to any particular use to be relatively elastic, for as the rent per acre that can be earned in it rises relative to that in other uses, land will be drawn from other uses that require land of similar quality and in similar situations.

THE SAVING PLAN

By *saving* we mean that part of the income that a household receives in any period which it does not plan to spend in that same period on the purchase of consumption goods and services. A household saves so that it may be able to buy consumption goods and services in some future period(s). Thus, if a household expects to receive an income of y_t in period t, and plans to spend c_t on current consumption, its planned saving for period t will be $y_t - c_t (= s_t)$. By *savings* we mean the wealth of a household at any point in time — that is, the total value at current market prices of all the material goods and claims to goods that the household owns. A household's savings may have been built up from its own saving in previous periods, or it may have been acquired by gift and bequest. The savings of a household are, then, the result of its own (or other households') saving. We may think of savings as constituting a reservoir that is being continuously augmented (or depleted) period by period by saving (or dis-saving). For each period, the household must make two plans: first, it must decide how much of its income to save, and second, it must decide how to hold its savings — whether in the form of money, securities or material goods. In this section, we shall describe both decisions, list and classify the data on which each depends, and show how each will be revised if any datum changes.

In making a saving plan, the household is deciding how much of the income that it expects to receive in the period lying ahead it will plan to spend on buying consumption goods during that period, and how much of it to reserve for buying consumption goods in future. The manner in which it will dispose of any

given income between consumption and saving will depend on the relative intensity of its desires for consumption now and consumption in the future, and on its estimate of its relative capacity to satisfy these desires. We shall illustrate the influence exercised by each of these by a simple example. Let us suppose that the household is making its plan at the beginning of period t, and that its planning horizon encompasses two consecutive periods — period t (the present) and period $t+1$ (the future). In Diagram 67, we measure quantities of present goods on the vertical axis, and quantities of future goods on the horizontal axis. Each point that lies between these axes will represent a combination of some quantity of present goods with some quantity of future goods, and each combination will promise a certain degree of satisfaction to the household. The household's preferences as between different combinations of present and future goods can be illustrated on the diagram by indifference curves. We have



OC – y_t/p_t CD – $y_{t+1}/1+i \times 1/p_t$

OA – y_{t+1}/p_{t+1} AB – $y_t(1+i)/p_{t+1}$

FC – s_t/p_t AE – $s_t(1+i)/p_{t+1}$

p_t – present period's price-level

p_{t+1} – next period's expected price-level

DIAGRAM 67

assumed that each indifference curve is convex to the origin,* because this accords with our knowledge of our own attitudes and because it is only from that assumption that we can deduce the kind of reactions to changes in incomes and interest rates that we observe in the real world. The indifference map in Diagram 67, then, shows us the household's tastes (as at the beginning of period t) for present and future goods and its preferences as between different combinations of them.

The ability of the household to obtain present and future goods will depend on the incomes it expects to receive and the prices it expects to have to pay for consumption goods and services in periods t and $t+1$ respectively, the rate of interest, and the stock of savings which it possesses at the beginning of period t. The way in which these limit its ability to satisfy its present and future desires can be shown in Diagram 67. We shall suppose that the expected incomes are y_t and y_{t+1} respectively, that the rate of interest is *i per cent* per period and that savings are initially zero. If the household were to eschew all present expenditures, the total monies at its disposal in period $t+1$ would consist of the income it expects to receive in that period plus the value of its current income and the interest it could earn on it by lending it, rather than spending it, during period t: that is, in period $t+1$ it would have $y_{t+1} + y_t(1+i)$ available for spending on consumption goods and services. Given the prices that the household expects to rule in period $t+1$, — p_{t+1} — we can calculate the quantity of goods that this money would buy: in Diagram 67, we assume that y_{t+1} (when spent in the manner described in Chapter 1) would buy OA goods ($=y_{t+1}/p_{t+1}$), and $y_t(1+i)$ would buy AB goods ($=y_t(1+i)/p_{t+1}$), so that the household's total command over $t+1$ goods if all its spending were done then would be represented by OB. If the household were to concentrate all its spending in period t, the sum of money at its disposal would consist of the income of period t — that is, y_t — and the 'present value' of the income it expects to

* Our assumption that each indifference curve is convex to the origin can be stated in another way. We shall define the marginal rate of substitution of future for present goods as the quantity of future goods the loss of which in the estimation of the household would just be compensated by an additional unit of present goods. The assumption that the curves are convex is, then, an assumption that the marginal rate of substitution of future for present goods decreases. An alternative name for this marginal rate of substitution is the marginal rate of time preference.

receive at the beginning of period $t+1$. If the household wants to spend next period's income now, it must borrow now from other households or firms, and pay interest on the loan until it can be repaid when the income of y_{t+1} accrues at the beginning of period $t+1$. The sum which it borrows must be such that the principal and the interest on it will be equal to y_{t+1} at that time — that is, it must be $y_{t+1}/(1+i)$.* Given the expected prices of consumption goods and services during period t, y_t will buy *OC* ($= y_t$ divided by p_t, which is an index of the prices of consumption goods in period t), and the present value of y_{t+1} will buy *CD* ($= y_{t+1}/p_t(1+i)$), so that if all the household's spending were done during period t, it could command *OD* present goods.

The straight line joining *D* and *B* passes through all the combinations of present and future goods that the household could enjoy with its expectations about its income, the relative prices of present and future goods and the rate of interest. Of these, it will prefer that denoted by *P*, where the line *DB* is tangential to one of the indifference curves: that is, it will plan to consume *OF* present goods and services during period t and *OE* goods during period $t+1$. During period t, it will save a sum of money (s_t) that would command *FC* present goods; in period $t+1$, it will spend y_{t+1} plus the then value of s_t, which will be $s_t(1+i)$ and which will command *AE* goods and services in that period. It is planning to save in period t and to dis-save in the subsequent period. In Chapter 1, we assumed (*supra*, pages 2–3) that the household had already decided upon its planned consumption expenditure per period; we have now shown how that choice is made.

The consumption-saving plan represented by the point *P* in Diagram 67 will be revised if there is any change in the household's tastes and preferences, its expectations about present and future incomes and prices, the rate of interest, or in its stock of savings. We shall illustrate the consequences of a change in any one of these. We must remember that this is more an illustrative than an explanatory exercise, for we have imputed the nature of the household's tastes and preferences as between present and

* If $y_{t+1}/(1+i)$ is borrowed at the beginning of period t, the sum that the household will owe at the beginning of period $t+1$ will be $y_{t+1}/(1+i)$ plus $i \cdot y_{t+1}/(1+i)$, which is $y_{t+1}/(1+i)$ all multiplied by $(1+i)$ or y_{t+1} — that is, the sum of money that the household will have available at the beginning of $t+1$ to repay the principal of the loan and pay the interest on it.

future goods from our knowledge of our own attitudes and our observations of its typical reactions to changes in the data listed above; our deductions cannot, therefore, contradict our experience. While one step backwards and one step forwards leaves us where we were, it also leaves us poised for another step forwards, for by assuming that the nature imputed to tastes and preferences from past experience will obtain in the future, we may hazard predictions about future behaviour.

First, the effects on the consumption-saving plan of a change in tastes and preferences. Let us suppose that a household plans to save more and spend less on consumption, while savings, incomes, prices and the interest rate remain unchanged. The cause of this revision in its plan must lie in some change in its tastes and preferences — it may have experienced some sudden psychological conversion that leads it to value future goods more highly than before. The new pattern of tastes and preferences is illustrated in Diagram 68: there, each indifference curve is relatively steeper near the Y-axis and relatively flatter as it approaches the X-axis, so that P_1 in Diagram 68 lies south and east of P in Diagram 67. Alternatively, we may say that the quantity of future goods that the household would surrender for each unit of present goods is now less than before. Conversely, if the household plans to save less, *ceteris paribus*, it must be because it has become less thrifty, and the implications of this change in its

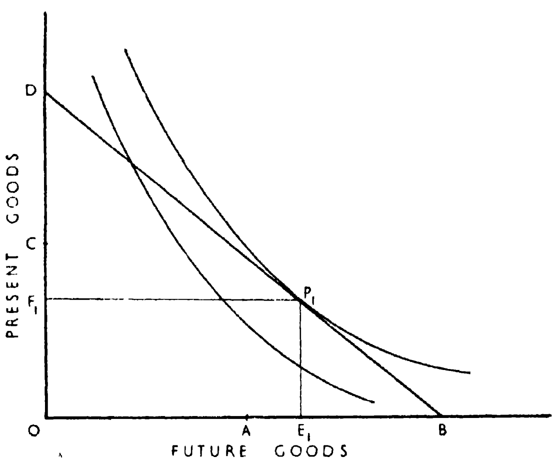


DIAGRAM 68

tastes and preferences can be seen by assuming that Diagram 68 illustrates the initial position and by comparing this with Diagram 67.

Second, the effects of a change in the rate of interest. In the figures in Diagram 69, all the combinations of present and future goods that the household could buy at the given expected prices and incomes and rate of interest of i per cent per period lie on the line BD. If the rate of interest should fall to i_1 per cent per period, *ceteris paribus*, this 'budget' line will move to B_1D_1, for now that the interest rate is lower, the command of the household over present goods will be greater and its command over

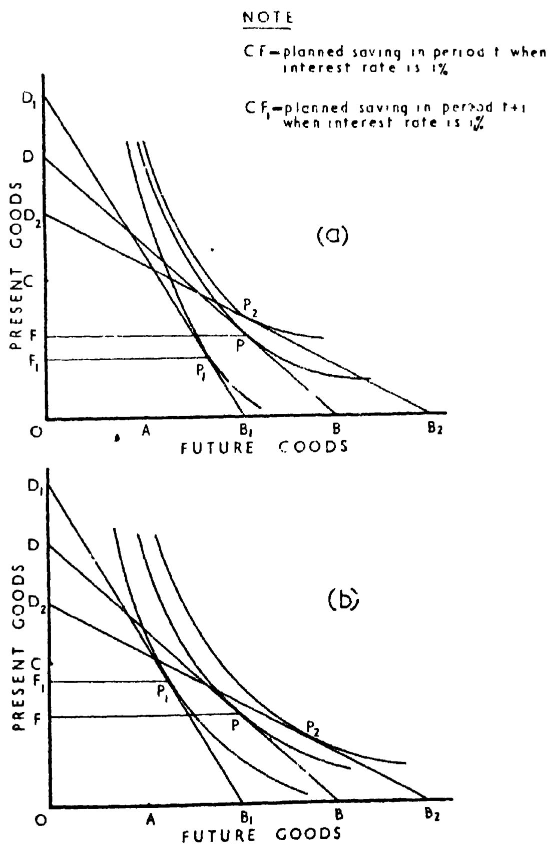


DIAGRAM 69

future goods will be less.* Of the combinations lying on B_1D_1, the household will choose P_1 where B_1D_1 touches an indifference curve. We know from experience that planned saving may fall or rise as a result of a relatively small fall in the interest rate — though the former reaction is the more common. By introspection, we can easily adduce reasons why either reaction may occur. If the lower rate of interest is expected to rule indefinitely, and if the household is intent on enjoying a given income from its savings over some span of future periods or on accumulating a given stock of savings at some future date, it may plan to save more now than before; and its tastes and preferences as between present and future goods would be as illustrated in Diagram 69(a). Typically, however, we would expect planned saving to fall when the rate of interest falls, for if all other things remain the same each sum of money that is set aside now will command a smaller quantity of goods in the next (or any future) period the lower is the rate of interest: this more usual reaction is shown in Diagram 69(b). Conversely, if the rate of interest were to rise to i_2 per cent per period, the 'budget line' would move to B_2D_2, and the new consumption-saving plan of the household would be denoted by P_2. At P_2, planned saving will generally be greater than at P, though it might be less. The relationships between planned saving

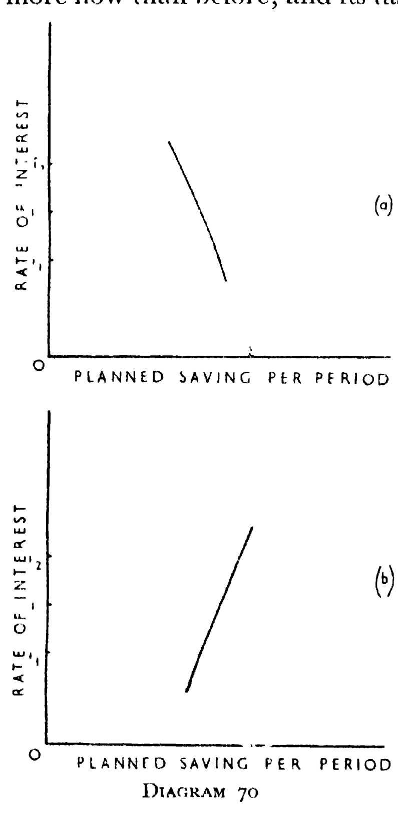


DIAGRAM 70

* The power to acquire present goods rises because the present value of next period's income (CD in Diagram 67) rises when the interest rate falls; command over future goods falls because the value in the next period of this period's income (AB in Diagram 67) falls when the interest rate falls.

and the rate of interest that are implicit in Diagrams 69(*a*) and 69(*b*) are shown explicitly in Diagrams 70(*a*) and 70(*b*) respectively. Irrespective of the direction of the change in planned saving, experience suggests that its magnitude is small: that is, in our jargon, the interest-elasticity of saving, while it may be positive or negative, will usually be 'low' -- that is, near to zero.

Third, the effect on the household's consumption-saving plan of a change in present and future incomes. Let us suppose that the household's income in period t rises, its tastes, the rate of interest, and its expectations about future income and present and future prices remaining the same. The consequences are illustrated in Diagram 71 by a movement of the 'budget line' from BD to B_1D_1,* and a change in the consumption-saving plan from that denoted by P to that denoted by P_1. The consequences of any other possible change in present income, all other things remaining the same, can be illustrated in a similar way. The relationship between present income and planned saving that is implicit in the points P, P_1, etc., in Diagram 71 is shown in Diagram 72. Most empirical studies have concluded that

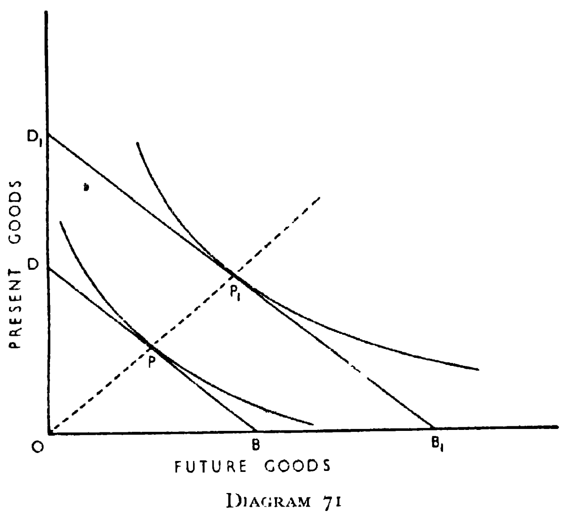


DIAGRAM 71

* It can easily be shown that B_1D_1 will be parallel to BD. The slope of BD is equal to $y_t + y_{t+1}/(1+i)/p_t$ divided by $(y_{t+1} + y_t(1+i))/p_{t+1}$ -- that is, to $p_{t+1}(1+i)/p_t$. The slope of B_1D_1, which is the 'budget line' when the present income has risen to y'_t is $(y'_t + y_{t+1}/(1+i))/p_t$ divided by $(y_{t+1} + y'_t(1+i))/p_{t+1}$, that is $p_{t+1}(1+i)/p_t$.

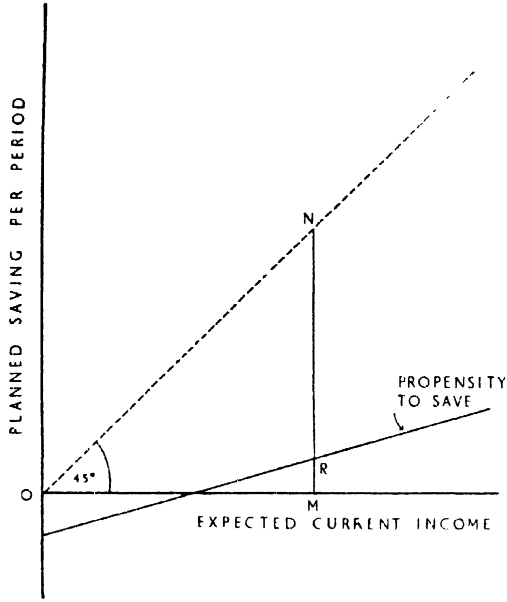


DIAGRAM 72

planned saving varies in the same direction, though not, of course, in the same proportion, as income, and the shape we have given to the indifference curves in Diagram 71 is such as to give this kind of behaviour. At very low incomes, saving may be negative — that is, there may be dis-saving — for the household might then be forced to realise its past savings, or, lacking these, to borrow money to supplement its meagre income; as its present income rises, saving soon becomes positive as it rises also. The relationship between the household's planned saving and its present income that is portrayed in Diagram 72 is called its *propensity to save.* Implicit in this is a relationship between planned consumption expenditure and present income. This latter can be shown on the diagram by drawing a line with a slope of 45 degrees through the origin: since the household plans to spend that part of each income that it does not plan to save, the planned consumption expenditure at each income will be equal to the vertical distance between the propensity to save line and the 45-degree line at that income.* Thus, if the present income were

* If we draw a 45-degree line through the origin in Diagram 73, we can see the relationship between planned saving and present income. Diagrams 72 and 73 are merely two different ways of saying the same thing.

OM ($=MN$), planned saving would be MR and planned consumption expenditure RN. The relationship between consumption and present income is shown explicitly in Diagram 73: this relationship is called the *propensity to consume.*

The propensity to save (consume) schedule that is graphed in Diagram 72 (73) shows us the amount that the household would plan to save (consume) at each level of present income, if its tastes and preferences, the rate of interest, its expectations about its future income and about present and future prices, all remained unchanged. The *proportion* of each income that the household would plan to save (consume), *ceteris paribus*, is called the *average* propensity to save (consume): in Diagram 72, the average propensity to save when present income is OM is equal to MR/OM, and the average propensity to consume at the same level of income is equal to RN/OM. As we have drawn the schedules, the former rises continuously and the latter falls continuously, as income rises, and this is in accord with the results of most empirical studies. The rate of change in planned saving (consumption) as present income changes is called the *marginal* propensity to save (consume) and it is measured by the slope of the propensity to save (consume) line over the appropriate range of income. In our diagrams, the propensities to save and

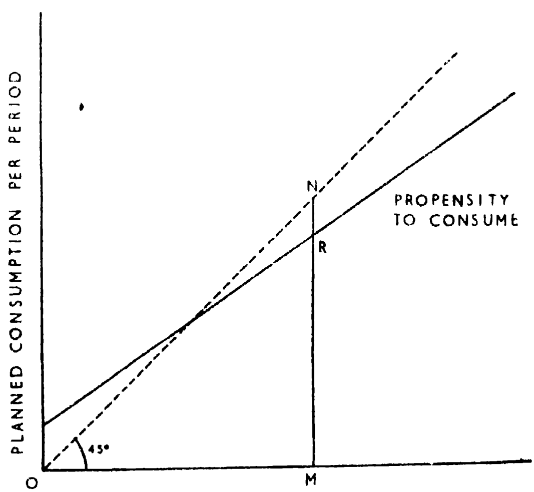


DIAGRAM 73

consume are drawn as straight lines, so that the marginal propensities to save and consume are the same at each income, and this also accords substantially with experience. From the manner in which we have measured them, it is clear that the marginal propensities to save and consume at each level of income when added together will be equal to unity.

In the same way, we may illustrate diagrammatically the manner in which the household's consumption-saving plan will be revised if its expectations about its future income changes, its present income and all the other data remaining the same. If it expects to receive a higher income in period $t+1$, B will move eastwards and D northwards, and the new consumption-saving plan will generally be such that planned saving in period t is less than before. Conversely, if the expected future income falls, the household will generally plan to save more now than before.

Fourth, the effects of a change in expected present and future prices. If present prices rise, then, *ceteris paribus*, the household's present money income, and the present value of its expected future income, will command a smaller quantity of present goods and services; the purchasing power of next period's income and of the then value of the present income will remain unchanged. The new 'budget line' will rise less steeply than the old one, and it is shown by BD_1 in Diagram 74. In the new consumption-

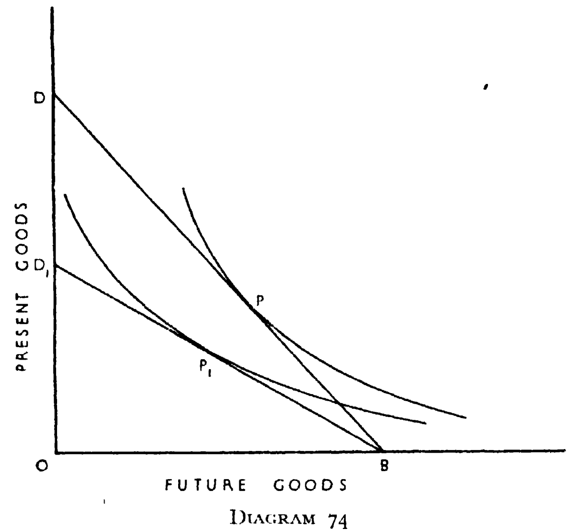


DIAGRAM 74

saving plan, which is denoted by P_1, the household is planning to save more now that present goods have become dearer relatively to future goods — its flow of consumption-spending is being redistributed over time in favour of those periods in which goods and services now appear to be relatively cheaper. Conversely, if present prices fall: the household will now plan to save less and to spend more now than before.

Lastly, the influence of savings on the consumption-saving plan of the household. We have assumed so far that the household has no savings when planning at the beginning of period t, and the indifference curves that we have drawn reflect its tastes and preferences for present and future goods in its knowledge that savings are zero. In our examples, the household will have accumulated savings by the beginning of period $t+1$. In describing its consumption-saving plans for periods $t+1$ and $t+2$, however, we cannot assume that its indifference map is unchanged, for the possession of savings will affect its relative valuations of present and future goods. Its present savings give it command over future goods, so that we would expect its desire to add to its savings — that is, to save during period $t+1$ — to be less intense. If we were to draw an indifference map that reflected the existence of savings, we would expect each curve in it to have the shape illustrated in Diagram 67 rather than the shape shown in Diagram 68. We shall not attempt to portray the influence of savings on tastes and preferences for present and future goods. Intuition, introspection and observation all suggest, however, that if its savings increase, the typical household will plan to save less and consume more of its present income, with its given expectations about its future incomes, present and future prices and the interest rate; and if the value of its savings declines, we would expect it, *ceteris paribus*, to plan to save more and spend less out of each level of present income. The influence of the value of savings on planned saving and consumption expenditure is called the 'Pigou Effect', and it is a force that tends to maintain full employment in macro-economic models in which all prices are flexible.* We shall return, though

* This volume is primarily concerned with micro-economics — that is, with the determination of relative prices, so that any treatment of macro-economic models and of the Pigou Effect lie beyond its scope. For the Pigou Effect, however, reference may be made to D. Patinkin, 'Price Flexibility and Full Employment', reprinted as Chapter 13 in *Readings in Monetary Theory*.

in no great detail, to the role of savings, for presently we shall be describing the forms in which a household may hold its savings and one 'form' in which savings may be held is consumption goods and services — that is, at any point in time the household may decide to spend all or a part of its savings.

We have assumed so far that only households save or have savings; in practice, however, firms do both. The income of a firm is its net revenue — that is, the sum of money that it expects to be left with after the fixed and variable costs of production have been paid out of the expected total revenue. A firm may either spend its income — that is, distribute it to the households in which the owners of the firm reside — or save it — that is, not spend it in this way. That part of the firm's income that is not distributed to the firm's owners is called 'business saving' or 'undistributed profits'. Our problem is to list the things on which the planned saving of the firm depends. If we regard the whole of the net revenue that the firm earns as accruing in the first instance to its household-owners, who then decide how much of it to spend and how much to save, then the preceding analysis might suffice: that part of the income that a household-owner receives which it does not plan to spend is left with the firm as undistributed profits. It may be tolerably realistic to view some firms in this light — especially firms that are owned by one or by a few people. Where the firm is a limited liability company, however, the nexus between ownership and control that this view assumes is much weaker. While the firm is owned by many households, it is controlled by its directors and managers, and the interests of these do not necessarily coincide. It is these latter who decide how much of the firm's net revenue will be distributed to its owners (to augment their incomes) and how much will be saved.

The determinants of the saving plan of the firm are similar to, though not the same as, those of the household's plan. We shall rest content with merely listing them. The analogue of the household's tastes and preferences as between present and future goods is the range of opportunities for earning net revenue that the firm expects to be open to it both now and in the future; we would expect the return per pound spent on buying inputs to decline in future as the firm planned to spend more and more pounds in the future and fewer and fewer pounds now, and vice

versa. Second, the firm's saving will depend on its expected present and future net revenues: if net revenue is expected to decline in future as compared with its present level, we would expect the firm to plan to save more now than it otherwise would. Third, the firm's plan will be influenced by the present relationship between input and output prices and how this is expected to behave in future: if input prices are expected to fall, *ceteris paribus,* we would expect the firm to save more now so that it will be in a better position to exploit the relatively cheaper inputs in the future. Fourth, the rate of interest: it is probable that business saving, like household saving, is relatively unresponsive to changes in the interest rate. Lastly, there are a number of influences that affect the firm's decision but which have no close counterpart with the household. A firm may desire to grow, and current saving is one method by which this objective may be achieved. Furthermore, firms may plan to borrow to meet future commitments or grasp future opportunities, and the strength of their desire to save now will tend to vary inversely with the ease and cheapness with which they expect to be able to procure money in the future when they need it. Given all these, we would expect business saving, like household saving, to vary in the same direction, though not necessarily in the same proportion, as current business income.

We have now described how the consumption-saving plan of the individual household (or the analogous plan of the individual firm) will be revised if there is any change in the household's (or the firm's) expectations about its present or future incomes, the prices it expects to have to pay for present and future goods, or in the rate of interest. From these revisions, two relationships are commonly derived, namely, the relationship between planned saving and the rate of interest, and that between planned saving and current income. We shall call the former the household's supply of saving, and we have called the latter its propensity to save. And our discussion in the previous pages has shown the direction in which each of these schedules will shift if any other planning datum should alter. The total supply curve of saving in each period may be obtained by adding together the planned saving of each household at each rate of interest: while some of the individual supply curves may have negative slopes and elasticities, it is unlikely that these will be

reflected in the shape of the total supply curve, for the majority of households will plan to save somewhat more as the interest rate rises. The total curve, like its components, will be interest-inelastic. It might be thought that the role that households play in determining the relationship between the rate of interest and other prices is summarised in the total supply curve of saving, in the same way as their role in determining the relative prices of other inputs is played by the total supply curves of them. This, however, is not so, and for two reasons. First, interest is the price received by those who lend money and paid by those who borrow it. This price is formally determined by the supply of, and the demand for, loans. The supply of loans — that is, of money for lending — is not, however, the same as the supply of saving, for saving is merely not spending on the purchase of current consumption goods and services: the money that is not so spent may be used in many ways, only one of which is to lend it. Second, as we shall see later (*infra*, Chapter 7) the relative prices of labour-service and of the services rendered by land are determined by the disposition of the total stock of labour and of land among their different uses. Saving, however, is not a stock: rather, it is a flow per period that augments the stock of savings. The flow of saving in any period is small as compared with the existing stock of savings; being small, we can neglect it as we did implicitly with the additions to the labour force (through, for example, a net excess of births over deaths) or to the stock of land (through reclamation, for example) earlier in this chapter. We shall see in the next chapter how households and firms in deciding upon the forms in which to hold their savings help to determine the rate of interest.

The propensity to save of all households and firms in an economy is a relationship between the economy's income (that is, the income of all firms and households) and planned saving. This cannot be obtained by simply 'adding together' the propensities to save of all the individual firms and households, and we can see why not by taking a simple example. We shall suppose that there are but two households in the economy and that their propensities to save are as drawn in Figures (*a*) and (*b*) in Diagram 75. Before we can calculate the total planned saving out of each level of the economy's income, we must know how the total income is distributed between the constituent house-

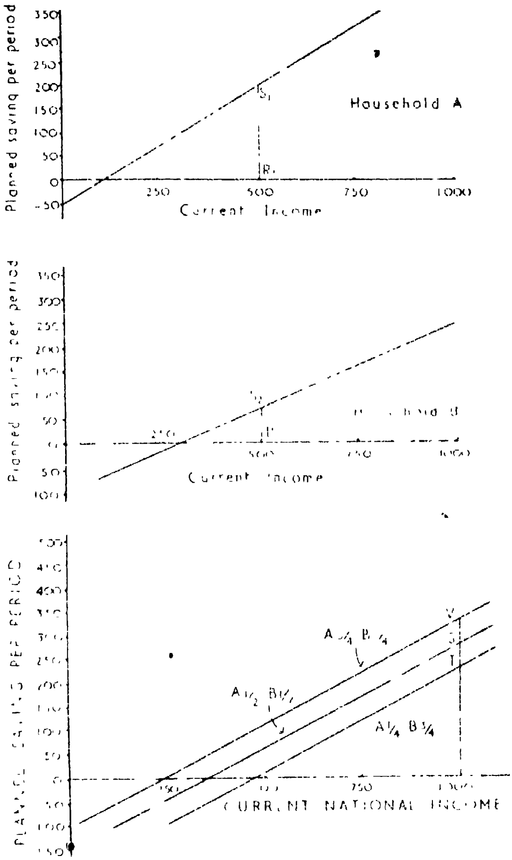


DIAGRAM 75

holds. Thus, if the total income is 1,000 per period, and if it is distributed equally between the two households, the total planned saving will be R_1S_1 plus R_2S_2 or RS in Figure (c). If the same income had been distributed in the proportions 1 : 3, with household A (which has the 'higher' propensity to save) receiving 250, and B (which would plan to save less out of each income than A) receiving 750 per period, the total planned saving would be less at RT. If the income had been distributed in the proportions 3 : 1, planned saving for the economy as a whole would have been greater at RV. This may be confirmed by supposing that A's propensity to save is given by the equation $s_a = -50 + \frac{1}{2} \cdot y_a$, and B's by the equation $s_b = -100 + \frac{1}{3} \cdot y_b$,

where s and y denote saving and income respectively, and the suffixes a and b particular households. If total income is 1,000, and if A and B each receive 500 per period, A will plan to save 200 and B 67, so that RS will be equal to 267. When A receives one-quarter and B three-quarters of this income, A will save 75 and B 150, so that RT is 225. When A gets three-quarters and B one-quarter, A saves 325 and B saves minus 17, so that RV is 308.

We could show in precisely the same way that the planned saving by both households at each level of total income will depend on how that income is distributed between them.* It is clear, then, that the propensity to save of the economy as a whole will rise, *ceteris paribus*, if income is redistributed in favour of households with relatively high propensities to save and that it will fall if the economy's income is redistributed in favour of households with relatively low propensities to save. When specifying the determinants of the propensity to save of an economy we must add to the list of influences that affect the saving decision of the individual household and firm the distribution of income between households and firms. The total propensity to save schedule shows the role that households play, and a part of the role that firms play, in determining the level of the economy's income. The manner and company in which these roles are played lie, however, beyond the scope of this volume.

In this section thus far we have concentrated on the saving decision; once this has been implemented, the household or firm must decide how it will hold the sum of money that it has saved until it requires it at some future time. We shall examine this savings decision in the next section.

THE SAVINGS PLAN

We shall suppose initially that at the beginning of period t, the household's (or the firm's) savings are equal to the sum of money that it has succeeded in saving from its income in the previous period. Its savings are then being held in the form of money, and it must now decide whether to hold them in that form during period t or to use the whole, or a part, of them to buy other things. The things other than money that a household

* Precisely the same problems arise in aggregating the propensity to consume schedules of the individual households.

or firm might acquire are bonds, bills, ordinary shares and physical goods.* At any moment of time, in any economy, there is a fixed stock of each of these, and the amounts by which these stocks are augmented (or depleted) over any interval of time such as a month or a year are insignificantly small when compared with the size of the stocks. In essence, a bond consists of a promise to pay some specific sum of money each year for a stated number of years with the further promise of a lump sum of money at the end of that period. These bonds may have been issued by central or local governments, or by industrial or commercial undertakings. An example of a bond is 2½ per cent Savings Bonds. These were issued 'at par' — that is, sold to households and firms at £100 each — by the British Government between May 16, and July 9. ; each Savings Bond consists of a promise to pay £2 10s. 0d. per annum, and the further promise of a lump sum of £100 some time between May 1, A bill consists of a promise to pay a sum of money at the end of a stated period, which is usually three or six months. An ordinary share is title of ownership to a share of a firm's assets and of the net revenue that these earn: thus, if a firm has issued one hundred ordinary shares, the holder of any one of them owns the one-hundredth part of the firm's buildings and machines, etc., and this entitles him to the one-hundredth part of the income they earn in each year. Lastly, a household or firm may buy physical goods, rather than ordinary shares which may be regarded as claims to physical goods and the net revenue they earn: the physical goods may be durable like land and buildings, or they may be storeable, like wheat, sugar and wool.

Now, at each point in time, each bill, bond, and ordinary share, and each kind of durable good or storeable commodity, has a price, and we shall suppose that these prices cannot be influenced by any individual household or firm. Thus, on October 12, the price of a 2½ per cent Savings Bond was £82 7s. 6d., and the price of a share in the Ford Motor Company was £2 3s. 9d.; on October 7, a Treasury Bill that promised £100 in 91 days' time could have been bought

* This list is not exhaustive. There are many hybrids that we have ignored, for bills shade into bonds, and bonds *via* preference shares shade into ordinary shares. It must also be remembered that bonds and ordinary shares are generic terms: for our present purposes, however, we are ignoring differences between one bond and another, and between one share and another.

for £98 18s. 8d. The problem that faces the household or firm is to decide how much of each of these to buy at their given and known prices with the money-savings at its disposal. This problem is formally the same as that which faces the household that has already decided upon its planned consumption expenditure and is considering what quantities of what goods to buy at their given prices. And its solution is similar: given its tastes and preferences, the household or firm will plan to hold that combination of bonds, shares, goods and money that promises to achieve whatever objective it is pursuing.

In Chapter 1, when describing the distribution of the planned consumption expenditure over the goods and services of every-day use, we assumed that the household's tastes and preferences were given. We made no attempt to list the qualities of these goods — such as colour, design, ability to satisfy hunger, and so on — on which tastes might depend; nor did we try to classify households — for example, into sybarites, aesthetes, sensualists, etc. — according to their attitudes towards goods and services. When discussing the disposition of savings, however, it is both usual and useful to do both these things. We shall first list the qualities (other than their prices) of bonds, shares and money respectively that would be relevant to a household or firm that was deciding whether or not to acquire them. Money may be used at any time to buy goods and services. It has the further advantage that its value in terms of money remains constant, for one hundred Bank of England notes will always be 'worth' £100 sterling. On the other hand, however, its real value — that is, the quantity of goods and services that a given sum of money can command — will vary inversely with the prices of those goods and services that the household or firm might wish to buy: if the prices of these goods should fall, then the real value or purchasing-power of the money that is held will rise, and vice versa. A government bond has the merit that it yields both a fixed annual income and a known sum of money at some time in the future. There are generally organised markets in which these bonds may be sold, so that the holder of a bond can always obtain money by selling it; he may, however, obtain more, or less, than he paid for it, for the price of a bond may fluctuate over time. For bonds that have been issued by private companies, we must add to this list the possibility of default —

that is, the chance that the company will not be able to honour its promises to pay a specified sum of money per year and a lump sum at some date in the future. With ordinary shares both the annual income (that is, the dividend) and the prices at which they may be bought and sold, might fluctuate. The dividend on the ordinary shares of a firm will tend to rise if there is a relative increase in the prices of the products that the firm produces and sells, and vice versa. Households and firms may store their wealth in the form of land and buildings: the prices of these may rise or fall as time passes and the income from them may fluctuate also. In these respects, they are similar to ordinary shares, and this should not surprise us, for the latter are in a sense merely claims to a particular share both in the ownership of a conglomeration of physical goods and in the annual incomes they earn.

The manner in which a household or firm will decide to hold its savings will depend (*a*) on its expectations about the future behaviour of the prices of bonds, of the prices and yields of ordinary shares and goods, and of the prices of the goods and services it might wish to buy with its savings at some time in the future, and (*b*) on its objective. A household or firm possessing savings may pursue any one of many aims: given its expectations about prices and yields, it may seek merely to maintain the money value or the real value of its savings, or it may seek to so use them that their expected money value is at a maximum at some future time;* alternatively, it may desire a stable money or real income from its savings, or it may wish either of these to be at a maximum. Since in this section our prime purpose is to show how households and firms help to determine interest rates as they decide how to hold their savings, we shall suppose that their objective is to maximise the expected money value of their savings. To justify this assumption at this stage, we must anticipate briefly: we shall presently see that interest rates vary inversely with the prices of bonds; bond prices vary in response to changes in the demand for or supply of bonds, and changes in the demand for bonds are mainly the result of changes in the plans of households and firms that are more concerned with the 'capital' value of their savings than with the income they yield. Given the objective, the precise combination of bonds, goods, shares and

* If the money value is at a maximum, the real value will be at a maximum also.

money that a household or firm will plan to hold will depend on its expectations of future prices and yields. Thus, if it expects bond prices to rise, we would expect it, *ceteris paribus*, to hold bonds rather than money or goods and claims to goods; if it believes that commodity prices are going to rise, we would expect it, *ceteris paribus*, to hold goods and claims to goods rather than money and bonds.

We shall describe the influence of expectations about yields and prices on the manner in which the individual household or firm decides to hold its savings with the help of a simple example. We shall suppose that the household or firm lives in a world where the only alternative to holding its savings in the form of bank notes and coins is to use them to buy one type of bond, namely, a perpetual bond. An example of such a bond is 2½ per cent Consolidated Stock ('Consols'), each unit of which promises a payment of £2 10s. 0d. per annum *in perpetuo*; there is no promise of a lump sum repayment. It can easily be shown that the rate of interest, which is the price paid for the loan of £100 for one year, will in this example vary inversely with the price of the bond: if a household buys a unit of 2½ per cent Consols for £50, it will receive an annual return of £2 10s. 0d. on the £50 spent on its purchase — that is, a return of 5 per cent per annum; if the price is £200, then the annual return of £2 10s. 0d. represents 1¼ per cent per annum.* Now let us suppose that the household has money-savings of £125 at the beginning of period t, that the price of a bond is then £62 10s. 0d. (that is, the current or market rate of interest is 4 per cent), that it expects the market prices of the goods and services it buys to remain unchanged (that is, that the real value of money to it will not alter), and that it is convinced that the bond price will be £50 at the beginning of period $t+1$, which is the next point in time when it will consider the disposition of its savings. Our problem is to

* Alternatively, we may regard the price of the bond as being the 'present value' (see *supra*, page 161) of the income stream that the bond promises: the present value (V) in our example is equal to:

$$\frac{2\frac{1}{2}}{1+i}+\frac{2\frac{1}{2}}{(1+i)^2}+\frac{2\frac{1}{2}}{(1+i)^3}+\dots+\frac{2\frac{1}{2}}{(1+i)^\infty}$$

where i is the rate of interest. This is a geometric progression, and its sum to infinity is $2\frac{1}{2}/i$. If V is 50, then i is $2\frac{1}{2}\times 100/50$, or 5 per cent; if V is 200, then i is $1\frac{1}{4}$ per cent.

describe the combination of money and bonds which if acquired now would have the maximum money value at the beginning of period $t+1$. In our example, it is clear that the household will plan to keep all its savings in the form of money, for by doing so their value at the beginning of the next period will be £125. If it bought one bond now and kept the remaining £62 10s. 0d. of its savings in money, the value of its wealth at the beginning of period $t+1$ would be £50 (the then bond price) plus £2 10s. 0d. received during period t plus the £62 10s. 0d. of money, or £115; if it had used all its savings to buy bonds now, the future value of its wealth would have been £105.

It is seldom, however, that a household or firm will be convinced that the bond price will be at some precise and unique level in the future, as we have assumed in the previous paragraph. In practice, the typical household or firm will be more or less uncertain about the future bond prices it may feel, for example, that there is little likelihood of the bond price exceeding 70 or falling below 60 at the beginning of period $t+1$, and that of all the likely prices lying between 60 and 70 some seem more likely than others. In these circumstances, for each present bond price, the household or firm can calculate what the value of its savings would be at the beginning of period $t+1$ for each present distribution of its savings between money and bonds on each hypothesis about the future bond price: thus, if its savings were now divided in the proportions 1 : 1* between money and bonds, the money value of this combination of money and bonds can be computed for each price at which bonds might be sold in the future, and this calculation repeated for each other possible distribution of present wealth. In our example, with a present bond price of £62 10s. 0d., money savings of £125 and the expectation that the price of bonds will be between 60 and 70, the money value of the savings at the beginning of period $t+1$ (ignoring the £2 10s. 0d. that is paid on each bond in each period) will lie between £122 10s. 0d. and £132 10s. 0d. if the present savings are apportioned equally between money and bonds, and between £120 and £140 if they are held wholly in bonds. In deciding how to hold its savings, the household must compare the range of probable values for its savings at time $t+1$ that would result from each distribution of savings between

* By value, bonds being valued at their existing price.

money and bonds with that resulting from each other distribution.

We know that households make choices like this, and that in the face of their uncertainty about the future level of the bond price they generally decide to distribute their present savings between money and bonds. We know little about the mental and emotional processes that precede such choices, but it is as if the household sought some compromise between the gain it would enjoy if the bond price should rise above its present level in future and the loss it would incur if the price of bonds should fall. Thus, in our example, with a present bond price of £62 10s. 0d. and an expected future price of somewhere between 60 and 70, we might think of the household as deliberating as follows: the probability of the bond price rising above £62 10s. 0d. makes bonds relatively attractive but the possibility of its being less than this makes them less attractive; it will compromise between these conflicting forces by holding some (probably the greater) part of its savings in the form of bonds to exploit the probability that the bond price will rise and some part in money to offset the possibility that the bond price will fall.

As always in economics, however, we are less interested in why the contents of a household's plan are precisely what they are than in how they would alter if any planning datum should change. If in our example the bond price were higher than £62 10s. 0d., then, *ceteris paribus*, we would expect the household to hold progressively fewer bonds and more and more money; for given the household's expectations, as the present bond price rises towards 70, the chances of it rising further become smaller and the chances of it falling become greater. If the bond price were 70 at the beginning of period t, the household may decide to hold all its savings in money, since it feels that there is little or no likelihood of the price exceeding 70 at the beginning of period $t+1$. If the bond price were now 60 — that is, at the level below which the household felt it unlikely to fall — the household might plan to hold all its savings in the form of bonds, since it believed that the bond price would most probably exceed 60 one period hence.

This relationship between the present price of bonds and the way in which the household would plan to apportion its savings

between bonds and money is illustrated in Diagram 76. Along the vertical axis, we measure the present bond price. The current or market rate of interest that is implied by each current bond price is measured along MM_1: thus, if the current price of $2\frac{1}{2}$ per cent Consols is 70, the market rate of interest is $2\frac{1}{2} \times 100/70$ or 3·57 per cent;* if the present price is 60, the current interest rate is $2\frac{1}{2} \times 100/60$ or 4·17 per cent. On the horizontal axes, through O and O_1, we measure the value of the household's savings in money; we have assumed that these are equal to OM ($= O_1M_1$) at the beginning of period t. If the bond price were 70, the household would plan to hold all its savings in money, for it feels that there is no possibility of the bond price being higher at $t+1$ than it is now. The line LO_1MN shows how the household will apportion the money savings that it possesses at the beginning of period t between money and 'money-used-to-buy-bonds' at each present bond price. From this diagram we can derive relationships between the present bond price (rate of interest) and the quantity of money or the number of bonds that the household would plan to hold, and these are illustrated in Diagrams 77(a) and (b) and 77(c) and (d) respectively. None of these dia-

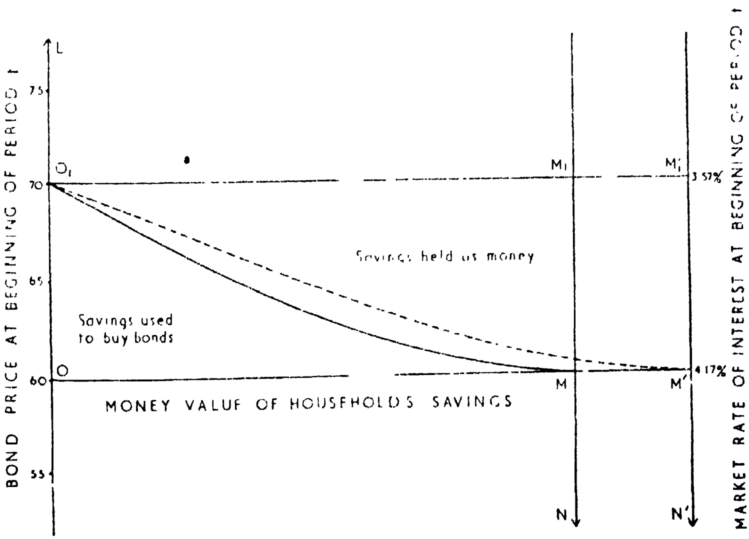


DIAGRAM 76

* See *supra*, page 200.

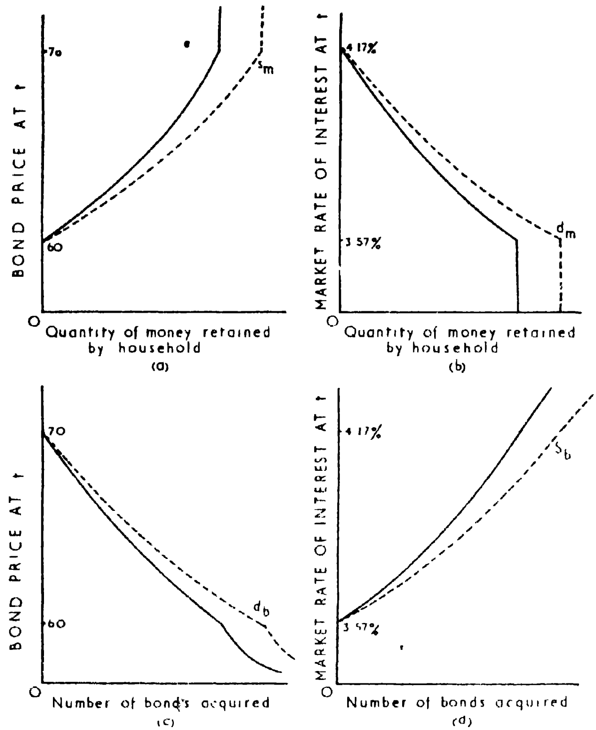


DIAGRAM 77

grams adds anything to Diagram 76. If we prefer curves which slope downwards to the right to be called demand curves, and those that slope upwards to the right to be called supply curves, we may label the relationship shown in Diagram 77(*a*) the household's supply curve of the service of holding money; that in Diagram 77(*b*) its demand for money; that in 77(*d*) its supply of the service of bond-holding, and that in 77(*c*) the household's demand for bonds. Since the sum of money that the household plans to spend on buying bonds rises as the present bond price falls from 70 to 60, it is clear that the demand curve for bonds is relatively elastic over this range of their prices. At current bond prices below 60, the demand curve for bonds will have unit elasticity, because the household will plan to spend the whole of its given money-savings on buying bonds. By the same token, the supply curve of the willingness to hold bonds will have unit elasticity at current rates of interest above 4·17 per cent.

The relationship shown in Diagram 76 between the present bond price and the apportionment of the household's wealth between money and bonds will change if the household revises its expectations about the range of values that the bond price is likely to have at time $t+1$. If, for example, the household felt that the bond price would then lie between 65 and 75, the line LO_1MN and the curves in Diagrams 77(*a*) and (*c*) would shift due northwards and those in Diagrams 77(*b*) and (*d*) would move due southwards, for at each present bond price the household would plan to hold less money and use more of its money-savings to buy bonds than before. And conversely if the household should feel that the bond price will assume a range of values lower than 60 to 70 at the beginning of period $t+1$.

If the value of the household's money-savings at time t had been greater than O_1M_1 at O_1M_1', then, *ceteris paribus*, the planned apportionment of savings between money and 'money-used-to-buy-bonds' would be shown by $LO_1M'N'$, and the curves in Diagrams 77(*a*), (*b*), (*c*), and (*d*) would have been in the positions shown by S_m, D_m, D_b and S_b respectively. And conversely if the household's money-savings at time t had been less than O_1M_1.

We have assumed so far that the household's savings are initially in the form of money and that it is deciding how much of them shall be used to buy bonds and how much shall remain in money. Our analysis needs little revision, however, if we suppose that the household, as a consequence of its decision at some time in the past, finds itself at the beginning of period t with its savings partly in money and partly in bonds. Let us suppose, for example, that at the beginning of period t, the household is holding x bonds and y pounds. If the bond price is then 70, the value of its existing wealth is £($70x$ plus y), and it will distribute this between money and bonds in the light of its expectations about what the future bond price will be at the beginning of period $t+1$, when it will again consider the disposition of its wealth. The relationship between the planned disposition of its savings and the current bond price could be illustrated by a diagram similar to Diagram 76, and from it we could derive relationships similar to those drawn in Diagram 77.

In our simple model economy, wherein each household and

firm must hold its savings either in money or bonds, we can derive for each household or firm the relationships shown in Diagram 77 between the current bond price or rate of interest and the amount of money or number of bonds that it will plan to hold. By simple summation, we can obtain relationships between the current bond price or interest rate and the amount of money or number of bonds that all households and firms in this economy would plan to hold. We can thus get a total demand curve for money as a store of wealth, or a total demand curve for bonds, or a total supply curve of the service (or willingness) to hold bonds, or a total supply curve of the willingness to store wealth in the form of money. At each instant of time, all these relationships are identical, and each of them summarises the role that households and firms play in determining the bond price or the rate of interest. We would expect these total curves to be relatively stable provided that the expectations of all households and firms about the future bond price are not revised in the same direction: thus, if all (or the majority of) households and firms feel that the future bond price will be higher than they had previously supposed, they will together plan to hold more bonds and less money at each current bond price than before, so that the demand for bonds (or the supply of the willingness to hold bonds) will move to the right, and the demand for money (or the supply of the willingness to hold savings in the form of money) will move to the left.

We have hitherto assumed that households and firms must hold their wealth either in the form of a homogeneous perpetual bond or in money, and we shall maintain this assumption when explaining the determination of the rate of interest in the next chapter. At this stage, however, we may indicate briefly how the analysis of the preceding pages may be extended to a world in which there are not only money and $2\frac{1}{2}$ per cent Consols, but also one kind of ordinary share. The price of ordinary shares, like the price of a bond, may fluctuate; but while the sum paid each year to the holder of each bond is fixed, that paid to the holder of each ordinary share is variable. At each point in time, when the household or firm is considering the disposition of its savings, it will know the current price of bonds and shares and it will have some expectations about the future behaviour of the bond price and of the price and dividend of ordinary shares.

These latter will usually vary in the same direction: thus, if each share entitles its owner to the one-ten-thousandth part of a firm's buildings and equipment, etc., and if the dividend paid per share should rise (say) from 2s. to 4s., the price that is paid for the share would generally rise also. A household's expectations of the direction in which the price of the ordinary share will vary will be related to the expected behaviour of prices: thus, if the household expects a relative rise in the prices of the products that a firm is making, it may expect the firm's net revenue and therefore the dividend it pays per share to rise. Given its expectations about the future behaviour of bond and share prices, the household will plan to hold that combination of money, bonds and shares that promises to achieve its objective. We would expect bonds to bulk the larger in its plan the lower is the present bond price, *ceteris paribus*. And vice versa. Similarly, given its expectations about the future behaviour of the bond price and of the price and dividend of ordinary shares, and given the present bond price, we would expect the household to use more of its savings to buy shares the lower the current share price, and vice versa. We may thus obtain a demand curve for bonds and a demand curve for ordinary shares: the former would shift if there were a change in the current price of ordinary shares, and the latter would move to a new position were there a change in the current bond price.

In this chapter, we have described in turn the sales plan of a household that sells labour-service, the services of land, and the services of the willingness to hold bonds or money. In each case, we have described the data on which each sales decision rests and the direction in which the sales plan would be revised should any datum alter. From the revisions that would follow a change in the present price of the service that the household is selling, *ceteris paribus*, we have derived a supply curve for each service. The total supply curve of carpenters' services, for example, isolates the role that the households in which there are carpenters or potential carpenters play in determining the relation between the hourly wage-rate of carpenters and that of other kinds of labour-service. Similarly, the total supply of wheat land summarises the role that the households that own it play in determining its relative price also. And the total supply of the

willingness to hold bonds summarises the role that the households who possess savings play in determining the bond price. In the next chapter, we shall describe how the relative prices of these services are determined. In doing so, we shall be combining the analyses of this and of the previous chapters.

CHAPTER 7

The Determination of the Relative Prices of Productive Services

In Chapter 5, we described the derivation of a firm's demand for any productive service that it might plan to buy; in Chapter 6, we derived the household's supply of any productive service that it might plan to sell. By aggregating these, we obtained the total or market demand and supply schedules respectively. The total demand for a productive service summarises the role that the firms that buy (or might buy) it play in determining its relative price as they implement their purchase plans. The price-determining role of the sellers of productive services is summarised in the total supply curve of each productive service. In this chapter, we shall describe how these roles are played, both in the short-run and in the long-run.

RELATIVE WAGE-RATES

The short-run supply curve of the services of carpenters, for example, is a schedule that shows us how the sales plans of carpenters would be revised if the only planning datum that altered were the expected hourly wage-rate: that is, it shows the number of hours of carpenters' services that all carpenters together would plan to sell in a given period of time at each price at which these services might be sold, *ceteris paribus*. The *cetera* that must remain *paria* are the number of carpenters, the tastes and preferences of each for real income and leisure, the prices of the goods and services that they might plan to buy, and their objective. The short-run demand for the services of carpenters is a schedule that shows us the number of hours of carpenters' services that firms would plan to buy in a given period of time at each hourly wage-rate, *ceteris paribus*. The *cetera* that must remain *paria* are the number of firms, the range of production possibilities open to each of them, the demand for the products that the firms are

planning to produce, the price of each other variable productive service that the firms are buying or which they might plan to buy, and the firms' objectives. The total demand and supply curves are graphed in Diagram 78: on the vertical axis, we measure the hourly wage-rate of carpenters, and on the horizontal axis, we measure the number of hours of carpenters' services that firms would plan to buy, or households plan to sell, in each period of time. The hourly wage-rate will tend towards the level *OW*, for only at that level will the purchase plans of firms and the sales plans of households for carpenters' services be consistent with one another.

The wage-rate will remain at *OW* per hour, with an even flow of sales and purchases each equal to *OQ* in each period, for so long as there is no change in the demand for carpenters' services or in the supply of them. The demand curve will shift to a new position, causing a change in the wage-rate in the same direction, if any one of the determinants of demand that are listed in the previous paragraph should alter; and we described in Chapter 5 (*supra* pages 146–8) how the demand would alter in response to a change in any one of these. The supply curve will shift if there is any alteration in any one of the determinants of supply, and we have already shown in Chapter 6 how supply will change when any one of these is altered.

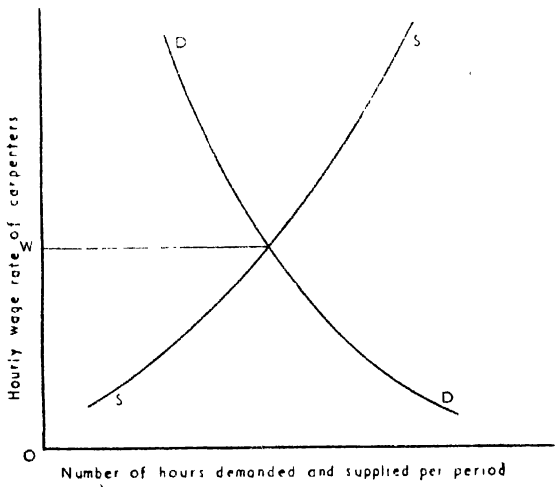


DIAGRAM 78

It must be emphasised that the preceding analysis, like that in Chapter 4, explains changes in the relationship between the hourly wage-rate of carpenters and the prices of products and of other productive services. Thus, if the preferences for leisure of carpenters become stronger, the supply curve in Diagram 78 will shift to the left, and the hourly wage-rate will rise as compared with (*a*) the prices of the goods and services of everyday consumption, and (*b*) the prices of other productive services.

In the long-run, as we have defined it for households as sellers (see *supra*, Chapter 6, pages 174–5), a household may change the kind of labour-service that it is selling: thus, in the long-run, a carpenter may renounce his skill and train as a bricklayer or bus-driver, or an agricultural labourer may become a carpenter. In the long-run, as we have defined it for firms as sellers and buyers (see *supra*, Chapter 2, pages 47–8, and Chapter 5, pages 155–6), a firm may change its method of production and so substitute carpenters for other productive services and vice versa. The influence of these long-run adjustments on the relative price of carpenters' services is illustrated in Diagram 79. The short-run demand and supply curves are represented by D_sD_s

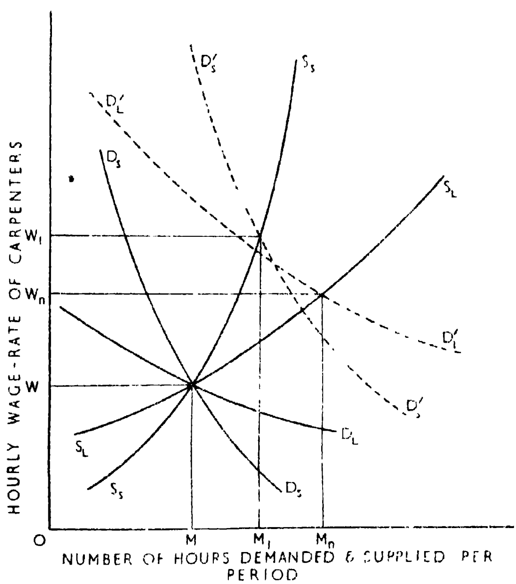


DIAGRAM 79

and S_sS_s respectively, and the long-run demand and supply curves by D_lD_l and S_lS_l respectively. As these curves are drawn, they portray a position of both long- and short-run equilibrium, for they all intersect at the hourly wage-rate OW. Let us now suppose that there is a permanent change in the preferences of households for the products that carpenters help to produce. This will cause an increase in the demand for carpenters' services, that is illustrated by rightward shifts in the short- and long-run demand curves from D_sD_s and D_lD_l to $D'_sD'_s$ and $D'_lD'_l$ respectively. In the ensuing short-period, the hourly wage-rate will rise to OW_1, and the planned purchases and sales of carpenters' services will rise from OM to OM_1. In the long-run, as households and firms revise their plans, the hourly wage-rate will decline towards OW_n, and the number of hours of work that are bought and sold will rise towards OM_n.

It is clear from Diagram 79 that the level towards which the hourly wage-rate will tend in the long-run will depend on the elasticity of the long-run demand and supply curves. For any given shift in the former, OW_n will be the nearer to OW the more elastic is S_lS_l, and vice versa. For any given shift in the long-run supply curve, OW_n will be the nearer to OW the more elastic is D_lD_l. The path by which the hourly wage-rate moves from OW to OW_n will depend on the expectations that each firm has about the price of its product and the hourly wage-rate which it expects to have to pay for carpenters, and on the expectations of each household about the future level of the carpenters' wage-rate when its members are contemplating a change in the kind of their labour-service. By making alternative assumptions about these expectations of firms and households, we may deduce a variety of paths by which the long-run equilibrium might be reached. These exercises are left to the reader, for they can be simply performed in the manner described in Chapter 4.

This explanation in terms of demand and supply analysis of the relationship between the hourly wage-rate of carpenters and the prices of products and other productive services has two main uses. First, it offers us a number of headings under which we may usefully and conveniently classify the causes of changes in the relationship between carpenters' wage-rates and other prices. The headings are what we have called the 'determinants' of the demand for, and of the supply of, carpenters' services.

Second, it helps us to predict the probable consequences of economic events on relative factor prices. It would be tedious to dwell upon the usefulness of the above analysis in diagnosing causes and exploring consequences, for there is little to add to what has already been said in Chapter 4 (see *supra*, pages 114–16). Rather, we shall rest content here with describing how demand and supply analysis may help us to interpret in a rather rough fashion the process of collective bargaining.*

Let us suppose that all existing carpenters are members of a trade union, and that all new entrants to the trade are eligible for membership. We shall suppose that the union's objective is to raise the hourly wage-rate. It may pursue this aim by restricting the supply (that is, by shifting S_lS_l to the left in Diagram 79),† by raising the demand (that is, by shifting D_lD_l to the right),‡ or simply by submitting a claim for a higher wage-rate to employers. If employers grant this claim to avoid a strike, the implications for the union may be illustrated in Diagram 80, in

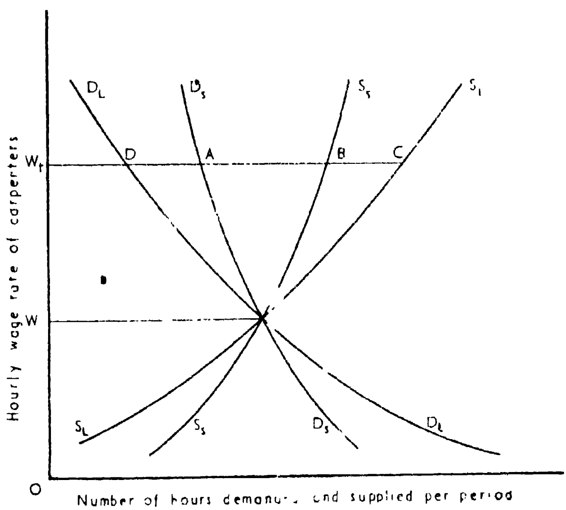


DIAGRAM 80

* It must be emphasised that demand and supply curves are very crude tools for the interpretation of the process of collective bargaining. More refined tools for elucidating this problem are described later in Chapter 11 under the heading 'Bilateral Monopoly'.

† This may be effected, for example, by limiting the number of hours for which each of its members might work per week.

‡ For example, by co-operating with employers in introducing new techniques of production.

which the curves have the same meaning as in Diagram 79. If the trade union threatens to strike unless the wage-rate is raised to OW_t, then the short-run supply curve of carpenters' services becomes W_tBS_s, and in each period there will be an 'excess supply' represented by AB — that is, there will be unemployment or underemployment of carpenters. The long-run supply curve will be W_tCS_l, and when long-run adjustments have been completed by employers, there will be an 'excess supply' of carpenters' services shown by DC. It is seldom, however, that employers capitulate so easily. Following the submission of the claim, there ensue negotiations, and the arguments by which the claim is supported and countered as these proceed can be roughly interpreted in terms of our demand and supply analysis.

Let us suppose that the existing wage-settlement was effected some time ago at t_o, and that the hourly wage-rate of OW that was fixed at that time was the then long-run equilibrium rate. This is a convenient simplifying assumption; whether or not it is true makes no difference to the substance of our argument. The short-run and long-run demand and supply curves at time t_o are shown by D_sD_s, D_lD_l, S_sS_s and S_lS_l respectively, in Diagram 81. Let us suppose that the trade union submits a claim for a wage-rate of OW_n at the beginning of period t_n. In supporting its

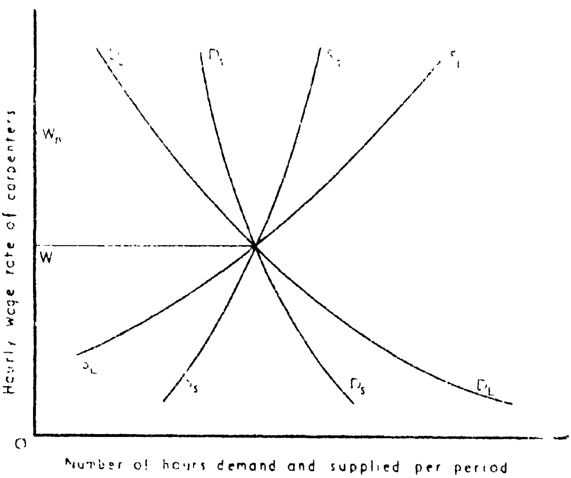


DIAGRAM 81

claim, the union may argue that the prices of the goods and services of everyday consumption (that is, the 'cost of living') have risen since time t_0. If this were the only change that had taken place since the last settlement, then the S_sS_s and S_lS_l curves would now lie in positions that are different from those they assumed at time t_0. The union may argue that the tastes and preferences of its members for leisure and real income are on balance such that the new curves lie to the left of the old ones. Secondly, the union may argue that the wage-rates in similar or comparable occupations have risen: if this alone had happened, then the long-run supply curve of carpenters' services would now lie to the left of its position at t_0. Third, the union may argue that the net revenues of the firms that employ its members have increased. The increase in net revenues may be attributed to a rise in the demand for the products that carpenters help to produce and/or to an increase in their physical productivities. In either case, the implication of this argument is that the short-run and long-run demand curves for carpenters now lie to the right of D_sD_s and D_lD_l respectively. The total effect of all such arguments is to suggest that the demand and supply curves at time t_n intersect at the wage-rate OW_n. The employers may deny the force of the union's contentions or question its estimate of the extent of the changes in prices, in wage-rates in alternative occupations, or in net revenues. In these ways, the employers may support their view that the long-run equilibrium wage-rate, appropriate to the conditions at time t_n, lies below OW_n. We are not here concerned, however, with the determination of the final outcome of the negotiations, for that requires a more refined analysis which will be described later in Chapter 11. At this stage, we wish merely to show how the arguments and counter-arguments may be interpreted in a rather crude way within the framework of demand and supply analysis.

In the preceding paragraphs, we have confined our attention to the determination of the relation between the price per hour of carpenters' services and the prices of all other products and services. The relative price of any other kind of labour-service, however, can be explained in precisely the same way. We shall next describe how the relative prices of durable goods, and of the services they render, are determined.

THE DETERMINATION OF THE RELATIVE PRICE OF A DURABLE GOOD

We have already derived the demand of an individual firm for a durable good like a weaving machine (see *supra*, Chapter 5, pages 159-63). This demand will only be operative when the firm is implementing its long-run plan. During any period, the total demand for the machine may be obtained by adding together (in the manner described in Chapter 5, pages 156-9) the demands of all the individual firms that are planning to buy it as they put their long-run plans into effect. This total demand for weaving machines is shown in Diagram 82 by the curve *DD*. The position and shape of this demand curve may vary from one period of time to another, depending on the number of firms that are deciding to implement their long-run plans. The short-run supply of machines may be obtained by the method described in Chapter 2 (see *supra*, pages 70-7), and their long-run supply curve by that described in Chapter 3 (see *supra*, pages 99-101). These are illustrated by S_sS_s and S_lS_l respectively in Diagram 82.

We shall suppose that the demand for weaving machines has remained stable at *DD* for long enough to enable the firms that

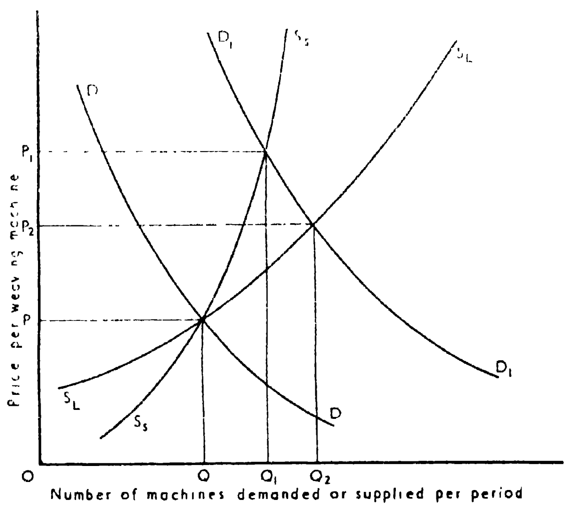


DIAGRAM 82

produce and sell them to make complete adjustments to it, so that the price per machine is initially OP. Let us now suppose that there is a permanent rise in demand to D_1D_1. In the ensuing short-run, the firms will expand their rate of output 'along' S_sS_s, and the price will rise to OP_1. Over the long-run, as the number and size of the firms that produce machines increases, the price will tend to fall to OP_2, and the number of machines that are being demanded and supplied in each period will tend to rise to OQ_2.

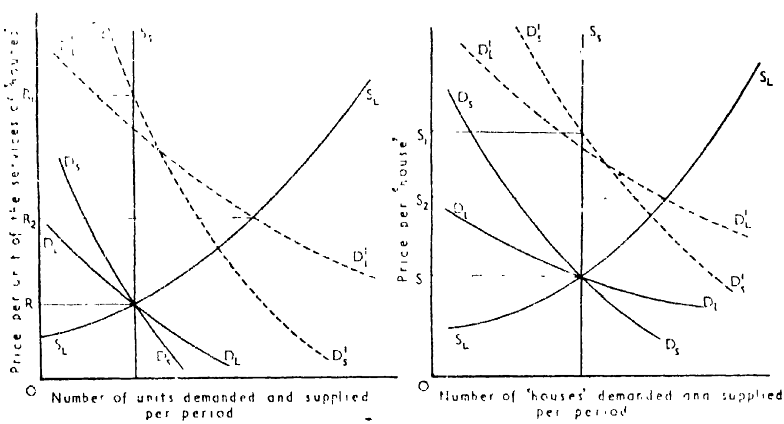
We know from Chapter 5 that the relationship that we have called the demand for a durable good depends, *inter alia*, on the conditions of demand for the product(s) that it helps to produce, the price of each other durable good and productive service in conjunction with which it may be used, and on the rate of interest, and that the relationship that we have called supply depends, *inter alia*, on the price of the productive services that are needed to produce it. If the rise in the demand for the machine in Diagram 82 is the consequence of a change in consumers' tastes and preferences for the product(s) it assists in producing, then the rise in the price of the machine from OP to OP_1 and its fall in the long-run to OP_2 represents changes in the relation between the price of the machine and the prices of products, other durable goods and productive services, and the rate of interest.

THE PRICING OF THE SERVICES OF DURABLE GOODS

The price that is paid for the services rendered by a durable good in each period is called rent in everyday usage. The durable good may be a dwelling house, a factory building, a plot of land or a machine. The explanation of the relative price of the services rendered by these is formally the same as our explanation of the wage-rate of carpenters in the first section of this chapter; we shall, therefore, deal with it briefly. If the shelter that dwelling-houses provide is bought by households, then the demand curve for it may be derived in the way described in Chapter 1, and we may, if we like, distinguish between the short-run and long-run demands for the services of dwelling-houses (see *supra*, Chapter 4, pages 133–6). If the services of the durable goods are being bought by firms to assist in the production of other goods and services, we may obtain the short-run and long-run demands for them in the manner described in Chapter 5 (see

supra, pages 139–59). In the short-run, the number of units of each durable good will be more or less fixed; there will, for example, be a given number of dwelling-houses available for renting, and of plots of land suitable in site and quality for the particular use we have in mind. The short-run supply curve of the services rendered by these will, therefore, be perfectly inelastic. In the long-run, the number of dwelling-houses may be depleted by dilapidation and by the use of dwelling-houses for other purposes, and it may be augmented by the building of new houses and the conversion of buildings that are now being used in other ways. The long-run supply curve of house-room will be more elastic than the short-run supply curve, and its elasticity will be the greater the greater is the elasticity of the long-run supply curve of buildings and the lower is the cost of converting dwelling-houses to other uses and of making buildings now used in other ways suitable for habitation. The same will be true of the long-run supply curve of land to a particular use: in the long-run land that is being put to other uses may be made suitable for the use in question and land now being used in this way may be made suitable for other uses.

The determination of the relative price of house-room, and its behaviour over the long-run, *ceteris paribus*, are illustrated in Diagram 83(*a*). We have supposed that at the rent per period of *OR*, there is initially both short-run and long-run equili-



(a) (b)

DIAGRAM 83

brium, and that this is upset by a permanent increase in the preferences for house-room. As a consequence, the short-run and long-run demand curves move to $D'_sD'_s$ and $D'_lD'_l$ respectively. The rent per period will rise to OR_1 and tend over the ensuing long-run towards OR_2. In Diagram 83(*b*), we have illustrated the determination of the relative price of the dwelling-house itself. In the initial equilibrium, the price is OS. If we ignore operating and maintenance costs, and if we are given the expected life of the house and the rate of interest, we know (see *supra*, Chapter 5, pages 160-3) that OR will be equal to $OS \,.\, i \,.\, (1+i)^n/(1+i)^n - 1$. When the demand for house-room rises, the derived demand for the dwelling-houses that provide this service will rise also: the price of each house will rise to OS_1 in the short-run and tend towards OS_2 in the long-run. The price of existing houses in the short-run (OS_1) and the rent per period of existing houses (OR_1) will be such that OR_1 is equal to $OS_1 \,.\, i \,.\, (1+i)^n/(1+i)^n - 1$. The price OS_1 exceeds the costs of building new houses; as these are provided, both the rent per period and the price per house will fall towards OR_2 and OS_2 respectively, and these must be such that OR_2 equals

$$OS_2 \,.\, i \,.\, (1+i)^n/(1+i)^n - 1.$$

At one time it was customary in economics to classify the factors of production into three groups—land, labour* and capital, and to call the price paid for the use of the productive agents that fall into each class rent, wages and interest respectively. This classification may be workable when an economy is in the initial stages of economic development, for individual factors of production may then fall easily into one or other of these groups. Further, it may then be useful, for each member of the economy may then own only factors that fall into a single group, so that this classification of productive factors and of the rewards paid to them may correspond fairly closely to the social classes landowners, red and green proletariat and capitalists re-

* Labour is sometimes divided into entrepreneurial and other labour, and the distinction between these is based on function. Entrepreneurship or enterprise may be defined as the labour which plans or co-ordinates the use of all other factors. It is sometimes defined as the factor which bears uncertainty—that is, whose reward reflects the existence of uncertainty. Since this is true of all factors (though to widely varying degrees) it may not be possible to identify uncertainty-bearing with a particular group of people, for all those who indulge in economic activity bear some uncertainty.

spectively. In a modern economy, however, what we ordinarily call land is land to which has been added capital and labour, and what is ordinarily called labour is human beings whose skills have been developed by education and training. In these circumstances, if we maintain the customary classification of factors, we must discard the classification of their rewards that accompanied it: for the price paid for the use of a plot of land whose quality has been improved by drainage and artificial fertilisers will then consist partly of rent (that is, the price paid for the use of land *per se*) and partly of interest; and the price paid for a doctor's services will be partly wages and partly interest also. Further, in modern social democracies, we do not find the same simple correlation between factoral groups and social classes, for fewer and fewer households now derive their incomes wholly from interest and with the diffusion of the ownership of capital goods, more and more households derive at least a part of their incomes from interest and dividends. Lastly, if a man decides to hold his money savings in the form of land, the price he receives by selling the use of the land will appear to him mainly as interest. In this chapter, our prime concern is to explain the determination of the relative prices of the things that firms buy. We have classified these things roughly into productive services, which make the whole of their contribution to production in the period in which they are bought, and durable goods, which yield their services over a succession of production periods. In calling the prices of some of these rents and wages, we follow ordinary usage. In explaining relative price behaviour, there is no need to try to break down the price of any particular service or durable good into the notional components appropriate to the three-fold classification of productive factors — that is, into 'rent', 'wages' and 'interest'.

A NOTE ON DIFFERENCES IN EFFICIENCY BETWEEN UNITS OF THE 'SAME' PRODUCTIVE SERVICE

In explaining the relative price of carpenters' services per hour in the first section of this chapter, we assumed that each carpenter *qua* carpenter was identical with each other. This meant that for employers an hour's work from any one carpenter was a perfect substitute for an hour's work from any other. In practice, however, carpenters may differ widely from one

another in efficiency: that is, all other things (such as the lay-out and organisation of the other productive services) being constant, an hour's work by carpenter *A* may yield a different output from that by *B*. In these circumstances, our explanation of the relative wage-rate must be modified.

Little modification is needed (*a*) if carpenters can be divided into sub-classes each containing carpenters that are of the same efficiency, or (*b*) if the labour-service that is being supplied by each carpenter can be reduced to a common denominator. If carpenters can be graded according to efficiency, and if within each grade there is a relatively large number of homogeneous carpenters, then the determination of the relative hourly wage-rate of each grade may be illustrated by a diagram similar to Diagram 79. The demand curves for the services of carpenters in each grade will be much more elastic than those shown in Diagram 79, for the services of grade I carpenters can now be substituted for those of the men in grades II and III, etc. If each carpenter differs from each other, so that sub-classification is impossible, it may be possible to reduce hours of work by heterogeneous carpenters to some common unit of efficiency. Thus suppose that the services of carpenter *A* are taken as the standard: if eight hours work per week from carpenter *B* are substituted for eight hours work per week from *A*, the relationship between the total outputs before and after the substitution will give a relationship between the efficiency of *A* and *B*, so that *B*'s services can be measured in the same units as *A*'s. If a controlled experiment of this kind were performed for all carpenters, their services might be reduced to the common efficiency unit, and this being done, the relative price per efficiency unit might be explained in the way illustrated by Diagram 79.*

If the contribution of each worker to output is identifiable and measurable, it may be possible to relate the wage to the output rather than to the time that is worked — that is, to have piece-rates rather than time-rates. When the price of labour-service is expressed as a simple piece-rate, the implied unit, in terms of which the work is being measured, is 'labour-service per unit of (homogeneous) output', and this provides a fair approximation

to the efficiency units that were described in the previous paragraph. Where there are simple piece-rates, the hourly wage-rate will vary between one carpenter and another roughly in proportion to their relative contributions to output—that is, to their marginal gross productivities. The less efficient worker, however, will tend to receive rather more than his marginal net productivity and the more efficient rather less: thus, if in one hour *A*'s output is four times that of *B*, *B*'s hourly wage will be one-quarter that of *A*, but since the quantities of the services of machines, plant and management required for each unit of *B*'s output are four times the quantities required for each unit of *A*'s output, *A*'s marginal net productivity will be more than four times greater than that of *B*.

If our purpose were to explain the precise wage that is being received by each carpenter, we would be forced to explore the implications of the fact that carpenters differ widely from one another in efficiency in much greater detail. We are primarily interested in this volume, however, in explaining changes in the relationship between the wages of carpenters and the prices of products and other productive services and factors. If the spread of efficiency among carpenters and the system of wage-payment are given, then our demand and supply analysis provides a useful framework within which to explain and interpret variations in the relative wages of carpenters. Thus, if the demand for the products that carpenters help to produce increases, then, *ceteris paribus*, we would expect the wage received by each carpenter to rise; if the wage-rate in other comparable occupations should fall, then, *ceteris paribus*, we would expect the average wage of carpenters to tend to fall.

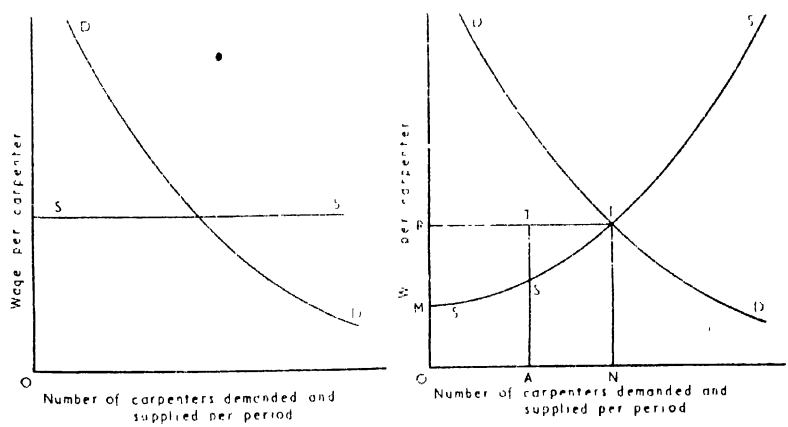
A NOTE ON 'ECONOMIC RENT'

Economic rent is the difference between the actual earnings of a unit of a factor of production and its supply price.* The actual earnings of a unit of a productive service or factor are the price that it receives for selling its services for a given period of time. Its supply price is the minimum sum of money that is required to retain it in its existing use. If the costs of transfer from

* 'Opportunity Cost' and 'Transfer Earnings' are synonyms for supply price. This definition of economic rent is substantially the same as that given by Joan Robinson, *Economics of Imperfect Competition*, Chapter 8, page 105.

one use to another are zero, then this will be equal to the maximum sum it could earn per period in any other use; if these costs are positive, its supply price will be equal to its highest earnings per period in an alternative use less one period's share of the transfer costs. Thus, if a carpenter can earn 210s. per week by working as a carpenter, and if the minimum sum that would induce him to do so is 185s. per week, his economic rent is 25s. per week.

If each carpenter is identical with each other *qua* carpenter, each will receive the same weekly wage, which will be determined by the demand for and the supply of carpenters' services. If all carpenters are identical not only *qua* carpenters but in all other respects also, then each will have the same supply price to carpentry, for the maximum earnings in alternative uses and the costs of transfer will be the same for each. In these circumstances, the long-run supply curve of carpenters' services will be perfectly elastic as in Diagram 84(*a*), and the weekly wage will be the same as the supply price of a week's work from each carpenter, so that no part of the earnings of any carpenter will be economic rent. If actual or potential carpenters are not equal in all other respects, then the long-run supply curve of carpenters' services will be less than perfectly elastic as in Diagram 84(*b*). They may differ from one another in that they are not equally versatile, so that the range of alternative occupa-



(a) (b)

DIAGRAM 84

tions open to them varies from one to another: the most remunerative alternative use for one might be driving a bus, for another acting as a waiter. They may have different attitudes towards the nature and conditions of the work in the various occupations open to them, and this by itself will mean that the wage that would induce *A* to become a carpenter might differ from that which *B* would demand. Lastly, the costs of transferring from other occupations to carpentry and vice versa might vary widely from one man to another: thus, to take the simplest example, *A* might live beside the firms which are demanding carpenters' services and *B* might live five miles away near a textile factory; other things being equal, *A*'s supply price to carpentry will be less than *B*'s. It is clear that economic rent will be a component of the earnings of most carpenters and the size of the component for any particular carpenter can be illustrated on Diagram 84(*b*). Thus, the economic rent received by the *OA*-th carpenter will be equal to *ST*, the difference between his actual earnings *OR* and his supply price *AS*; that of the *ON*-th carpenter will be zero for the weekly earnings of *OR* are just sufficient to induce him to acquire or retain this skill. The area *RML* shows that part of the total earnings of all carpenters who are at work (*ONLR*) that is economic rent.

It is clear from Diagram 84(*b*) that economic rent will be a more important constituent of the actual earnings of a productive service the less elastic is its supply curve. The elasticity of its supply curve will depend, *ceteris paribus*, (*a*) on the manner in which we define a productive service and the uses to which it might be put, and (*b*) on time. If we adopt broad definitions of productive service and use, economic rent will be the greater, and vice versa. Thus, if we group all the natural, non-human agents of production together and call them land, and if we define agriculture as the sole use of land, then the supply curve of land to agriculture will be perfectly inelastic, for on these definitions land is quite specific to agriculture so that its supply price is zero: the whole of the actual earnings of landowners will, therefore, be economic rent. Given the definition of productive service, economic rent will be the less the narrower are our definitions of use: thus, if we distinguish between the use of land for growing corn, wheat, potatoes, apples, and so on, the supply price of land to each of these uses will be positive, so that

only a part of its actual earnings in any particular use may be called economic rent. If we go further and define as a separate use the growing of potatoes for Mr. Smith, who is one of many growers of potatoes in a locality, then the supply of land to him will be perfectly elastic, for the supply price of land to him will be its market price.

On any set of definitions of productive service and use, the supply price of a unit of a productive service to any use will depend on the range of alternative uses that is open to its owner, and this tends to vary directly with time. If we define the short-run as a period within which a unit of a productive service cannot move from one use to another, then its supply price to its present use will be zero, and the whole of its current earnings will be economic rent. In the long-run, units of a productive service may move from one use to another: thus, carpenters may become bricklayers, and a firm whose past savings are embodied in a weaving machine may realise these from depreciation allowances and use them to buy a spinning machine or a tractor. In the long-run, therefore, the supply price of each unit of a factor to its existing use will be what it could earn in its next most remunerative use (if we ignore transfer costs), and not the whole of its actual earnings will be economic rent. Economic rents that appear in the short-run have been called quasi-rents* to draw attention to the fact that in whole or in part they are likely to be temporary. It would seem, then, that wide definitions of productive services and use have the same effect on the size of the economic rent as a narrow planning horizon, and that narrow definitions have similar consequences to a lengthening of the planning horizon.

As we have defined economic rent, it is a surplus: if, when the price of each product and productive service were at its long-run equilibrium level, all economic rents were appropriated by the state, no unit of any productive service would have any incentive to change its use, for post-tax earnings of each would be equal to its supply price to its existing occupation. For this reason, the notion of economic rent has been of some importance in the history of public finance: with its aid, it was possible to conceive of a system of taxation that would not directly affect

the pattern of resource-use within an economy. Even if such a tax system were practicable, however, it would alter the distribution of income between the owners of productive services. Since tastes and preferences differ from one person to another, there would be a change in the pattern of demand, and hence in the pattern of relative product and factor prices, and this in turn would cause a change in the pattern of resource-use in the economy.

THE RATE OF INTEREST

By a rate of interest we mean the price per unit that is paid for a loan of money for a period of time. Conventionally, the unit of money is £100, and the period of time is one year, so that the price is usually expressed as a rate *per centum* per annum. In this section, we shall attempt to explain how this price is determined. In doing so, we shall simplify heroically by assuming that all loans are riskless, that all are made for an infinitely long period of time, and that borrowers borrow by selling irredeemable bonds and lenders lend by buying them. We shall suppose also that the number of bonds that is currently issued in any period of time is insignificantly small as compared with the quantity of bonds that have already been issued. On these assumptions, all bonds will be homogeneous, the market price of bonds will be determined primarily by the demand for the existing stock of bonds, and the relationship between the market price of bonds and the 'nominal rate of interest' on them will give us the current or market rate of interest.* Later, we shall modify our explanation by assuming that bonds are not homogeneous, either because they are issued for different periods of time, or because there is uncertainty about whether borrowers will meet their promises to pay the nominal rate of interest on their bonds and repay the principals.

We have already seen that the manner in which a household will plan to distribute its savings between money and bonds during the period that lies ahead depends on the current bond price (rate of interest), the household's expectations about the future level of the bond price and its objective. In Chapter 6,

* If each bond promises a payment of £2 10s. 0d. per annum indefinitely, we shall call the £2 10s. 0d. the 'nominal rate of interest' on the bond. If the market price of the bond is £50, then the market rate of interest will be 5 per cent (see *supra*, Chapter 6).

we have shown what the planned disposition of savings would be at each current bond price, assuming that the household's expectations and objective remained unchanged. From this relationship, we derived the household's supply of the willingness to hold money, its demand for money as a store of wealth, its demand for bonds and its supply curve of the willingness to hold bonds.* By adding together the demand curves for bonds of all the households and firms in the economy, we obtain the total demand for bonds. This shows us the number of bonds that the firms and households would plan to buy at each current bond price, given their objectives, their expectations about the future level of the bond price (market rate of interest), the initial distribution of the savings of each household and firm between money and bonds, and the initial distribution of the existing stocks of money and bonds between households and firms. These last two determinants of the demand for bonds mean, in effect, that both the number of bonds and the quantity of money in the form of which savings might be held, must be assumed constant. At each point in time, there will be a given stock of bonds: that is, their supply curve will be totally inelastic. These demand and supply curves are graphed in Diagram 85, where we measure the current price of bonds on the vertical axis and the number of bonds demanded and supplied on the horizontal axis. The mar-

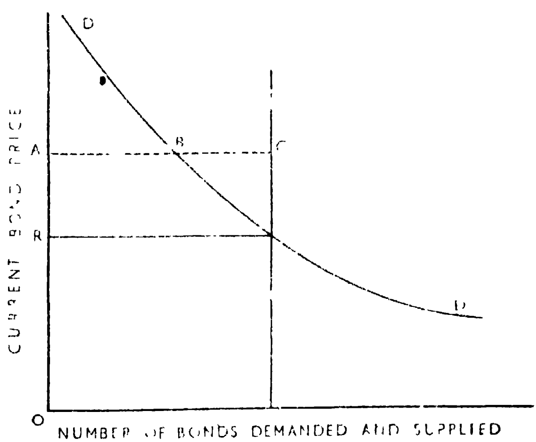


DIAGRAM 85

* See Diagrams 77(*a*), (*b*), (*c*) and (*d*) respectively.

ket price of bonds will tend towards the level OR, for only at that price will the number of bonds that firms and households plan to purchase be equal to the number of bonds that are available for purchase: that is, only at the price OR will the public be willing to hold the existing stock of bonds. If the bond price is OR, and if the nominal rate of interest is $2\frac{1}{2}$ per cent then the market rate of interest will be $2\frac{1}{2}/OR$.

It can be seen from the diagram that households and firms will only be willing to hold the existing stock of bonds when the current bond price is OR. Thus, if the bond price were now OA, the members of the economy as a whole would be holding BC more bonds than they wished to hold at that price: those holding more bonds than they desire would attempt to sell them, and the pressure to sell bonds would lower their price. As the bond price fell, the pressure to sell bonds would diminish and the inducement to buy them would rise. And conversely, if the current bond price were less than OR: the 'excess demand' for bonds will *ceteris paribus* raise their price to OR. If all firms that borrow money do so by selling bonds that are identical with those already in existence, and if the flow of new bonds in any period of time is insignificantly small when compared with the stock of bonds already issued, then the sales of new bonds will not affect the bond price: that is, the market rate of interest of $2\frac{1}{2}/OR$ will be that at which new loans can be obtained.

The bond price will move to a new level if the demand for bonds alters. If, on balance, firms and households expect that the future bond price will be higher than they had previously supposed, then the demand for bonds will increase and the current bond price will tend to rise; and conversely. If there is an increase in the quantity of money that is available for holding as an asset, then, *ceteris paribus*, the demand for bonds will rise; for the demand for bonds of each individual firm and household that receives a part of the increase in the quantity of money will swivel rightwards (see *supra*, Chapter 6, page 205) so that the total demand will rise also. The increase in the quantity of money available for use as a store of value might be a consequence of a redistribution of an existing stock of money between this and other uses (see *infra*, pages 229–31), or of an increase in the total stock of money. In most modern economies, the increase in the total stock of money is effected by the purchase of

bonds by the monetary authorities, and the stock of money is depleted by the sale of bonds. The purchases and sales of bonds by the monetary authorities with the aim of changing the quantity of money are called 'open-market' operations. If the demand curve in Diagram 85 is defined as the total of the demands for bonds by firms and households (that is, the public) and by the monetary authorities, and if the quantity of money is increased by bond purchases by the latter, then to the increased demand for bonds by the public as a consequence of the increased quantity of money we must add the demand for bonds by the monetary authorities. In these circumstances, the bond price will rise by more than it would have risen if the quantity of money had been increased by other means. Alternatively, if the demand curve in Diagram 85 is defined as the total demand by the public for bonds, then the effect of bond purchases by the monetary authorities will be illustrated by a leftward shift in the supply curve of bonds, for now that more bonds are held by the authorities fewer will be available to the public.

The explanation of the determination of the current bond price (rate of interest) may be presented in terms of the demand for and supply of money. From the manner in which the individual would plan to revise the disposition of his savings between money and bonds in response to changes in the current bond price, we can derive his demand curve for money as a store of value (see *supra*, Chapter 6, pages 201–4, and Diagram 77(*b*)). When the individual demand curves of households and firms are summed together, we obtain the total demand for money as a store of wealth. This shows us the number of units of money that the firms and households in the economy would plan to hold at each current rate of interest, given their objectives, their expectations about the future level of the bond price (rate of interest), and the economy's stocks of money and bonds.

At any point in time, there will be a given quantity of money in an economy. The whole of this, however, will not be available to function as a store of value, for some part of it must act as a medium of exchange. We have already described the purchase and sales plans of households: when the sales plan is implemented, goods and services are exchanged for the money that constitutes the household's income; when the purchase plan is implemented, the sum of money that we called the planned

consumption expenditure is exchanged for goods and services. Since the consumption expenditure is mainly financed from income, money is here acting as a medium through which the productive services that the household owns are exchanged for the goods and services that it wants. If each household received payment for what it sells at the same moment as it pays for what it buys, it would require no stock of money to finance this exchange. Typically, however, incomes are received at discrete intervals, while consumption spending takes place more or less continuously, so that at each instant of time, a household will have some sum of money designed for spending which is as yet unspent. Given the pattern of spending, this sum will be the greater the larger is the household's income and the less frequently is it paid. Thus, if a household receives £8 on Friday evening in payment for the services sold during the previous seven days, and sets aside £7 for consumption spending at an even rate of £1 per day during the seven days that follow, its average daily stock of money-for-spending will be £3.* If the same income were paid fortnightly, then on the same assumptions the average daily amount of money held as a medium of exchange would be £6 10s. 0d. If the weekly income had been £16, planned spending £14 and daily expenditure £2, then, *ceteris paribus*, the average daily holding of money would have been £6. The amount of money that a household holds to bridge the gap between the receipt of income and its expenditure is called its transactions balance.

For each firm in an economy, money acts as a medium through which its flow of products is exchanged for the flow of productive services needed to make them. Since the productive services are used to make the firm's products, payment for the former may (and generally does) precede the receipts of money from the sale of the latter. Given the customary intervals at which the firm pays for the things it buys and receives payment for the things that it sells, it will require some sum of money to bridge the gap between its payments and receipts. This sum is called its 'working capital' or transactions balance. Given the relationship between the frequency of receipts from sales and

* Assuming that the spending is done first thing each morning, its stock of money on Saturday will be £6, on Sunday £5, and £4, £3, £2, £1 and £0 on Monday, Tuesday, Wednesday, Thursday and Friday respectively. The average daily stock will be the sum of these divided by 7 — that is, 21/7 or 3.

the frequency of its expenditures on the purchases of productive services, the size of a firm's transactions balance will be the greater the greater are its receipts. The receipts of all the firms in an economy will depend largely on the level of spending by all the households, and that in turn will depend on the aggregate income of the households. That part of the total quantity of money that is required to facilitate the current transactions of households and firms will, therefore, depend mainly on the level of the economy's income.

If we are given M, the number of units of money available for all uses in an economy, and if we are given the quantity (M_1) that is required for the transactions balances, then $M - M_1$ or M_2, will be the number of units available to satisfy the demand for money as a store of value. If M is assumed given, and if we suppose that the transactions balances will not vary with any likely change in the rate of interest, we may conclude that M_2 will be inelastic with respect to the rate of interest over the range in which it is likely to cut the demand curve for money. In Diagram 86 we measure the market rate of interest on the vertical axis, and the quantity of money demanded and supplied for use as a store of value on the horizontal axis, *DD* is the demand for money and *SS* the supply curve of it, and for simplicity's sake the latter is drawn as being perfectly inelastic. The market rate of interest will be *Oi*, for only at that level will that

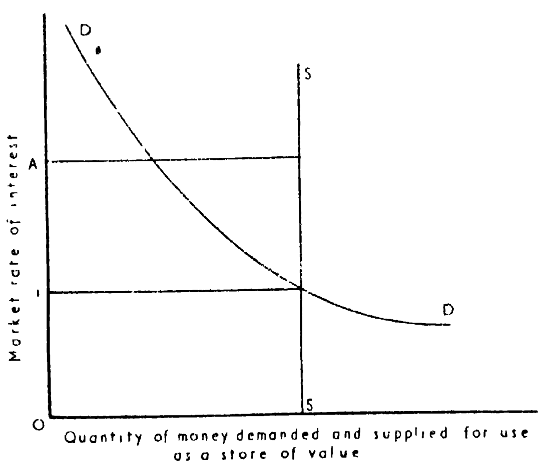


DIAGRAM 86

part of their savings that the public wish to hold in the form of money be equal to the quantity of money that is available for acting as a store of wealth. If the market rate of interest were at OA, then firms and households taken together would find themselves holding a larger part of their savings in money than they desire. This would impel them to reduce their holdings of money by buying bonds, so that the bond price would tend to rise and the market rate of interest to fall. The desire to reduce their money holdings would persist until the rate of interest had fallen to Oi. The market rate of interest Oi corresponds to the bond price of OR in Diagram 85.

The rate of interest will move to a new level if there is any change in the demand for money as a store of wealth — for brevity's sake, we shall follow common usage and call this the *speculative* demand for money — or in the quantity of money available for meeting this demand. Thus, if the public on balance expect the rate of interest to be higher in the future than they had previously supposed, the speculative demand for money will increase, and the market rate of interest will rise. If the level of income should rise, then M_1 will rise, and if M remains the same, M_2 must fall, and, *ceteris paribus*, the rate of interest will rise. If, while the economy's income remains unchanged, M is reduced by the sale of bonds by the monetary authorities — that is, by open-market operations — then, *ceteris paribus*, the rate of interest will rise. If the DD-curve in Diagram 87 shows the demand of households, firms and the monetary authorities for money, then these open-market operations will shift the demand curve for money to the right through a horizontal distance equal to the value of the bond sales; if SS represents the initial supply of money for speculative uses, it will shift leftwards to S_1S_1 as a consequence of the open-market operations. In these circumstances, it can be seen that the market rate of interest will rise to Or: that is, when M_2 is reduced by bond sales by the monetary authorities, the rate of interest will rise by more than it would have risen had the same reduction in M_2 been effected without open-market operations.*

In this section thus far we have concentrated on explaining the determination of the market rate of interest. Let us now suppose that at the beginning of some period t, there is a permanent

* Compare *supra*, pages 228–9.

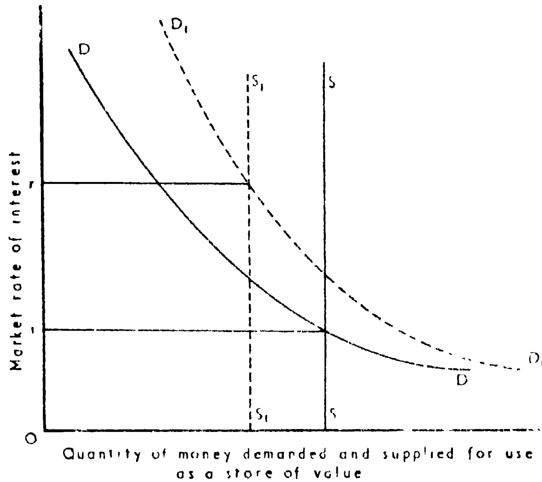


DIAGRAM 87

rise in the speculative demand for money. As we have already seen, the interest rate will rise: but will the interest rate remain stable thereafter at its new and higher level, or will the new interest rate cause changes that will in their turn tend to move it towards some 'long-run' equilibrium level? It will be recalled that similar questions were asked in Chapter 4 and in the earlier sections of this chapter: we have seen that if there were a permanent rise in the demand for, say, butter, its price will rise in the short-run; this will lead firms to revise their long-run sales and purchase plans and as these are implemented the price of butter will tend to fall to some long-run equilibrium. The 'long-run' behaviour of the interest rate lies rather outside the limits of this volume. We shall nevertheless offer a brief sketch of one way in which we may seek to explain it; for a fuller description of the relationships that we shall use, the reader is referred to any text on macro-economics.*

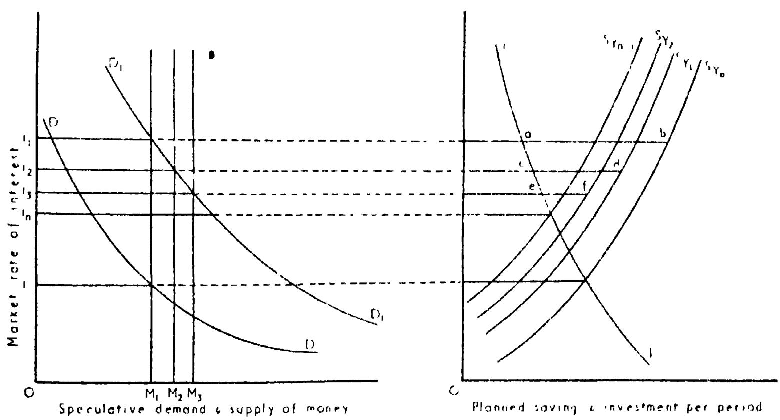
We shall define the 'long-run' equilibrium rate of interest as that rate at which the economy's income will remain stable from one period to another: thus, if we denote total income by Y, and successive time periods by the subscripts t, $t+1$, $t+2$, ... $t+n$, when the rate of interest is at its 'long-run' equilibrium level, Y_t will be equal to Y_{t+1}, and Y_{t+1} to Y_{t+2}, and so on. By the

economy's total income we mean the value at current market prices of all the productive services sold by households within a period plus the net revenues earned by firms in that period. We shall define a period as the length of time required for expenditures by households and firms on the purchase of currently produced goods and services to generate income. The income-generating expenditures within each period may be roughly classified into expenditures on currently produced consumption goods and services, which we shall call consumption and denote by C, and expenditures on newly produced investment goods, which we shall call investment and denote by I. Within any period t, then, on these definitions: $Y_t = C_t + I_t$. We have already seen that the level of consumption spending and of saving (that is, of not spending on consumption) depend, *inter alia*, on income (*supra*, Chapter 6, pages 187–91 and 194–6), and for our present purposes we shall suppose that planned consumption and saving for any period depend upon, and together exhaust, the previous period's income: that is, $C_t + S_t = Y_{t-1}$. If $Y_{t-1} = Y_t$, then $S_t = I_t$. When the interest rate is at its 'long-run' equilibrium level, on our definitions, then in each period planned saving must be equal to planned investment expenditure.

In Chapter 6, we derived a saving supply schedule for an economy: this was a relationship between the rate of interest and planned saving, given the tastes and preferences for present and future goods, current and expected future incomes and prices, and the distribution of income. In Chapter 5, we described the purchase plan of a firm for an investment good: the number of units of any investment good (such as a machine) that the firm will plan to buy will depend on its price, the firm's knowledge of productive techniques, the price of each other investment good and productive service, and the rate of interest. And we saw that the number of machines that the firm would plan to buy in any period would vary inversely, *ceteris paribus*, with the rate of interest. If we suppose that the price of all goods and services are constant (as they would be if the total supply curve of each of them was perfectly elastic), we may obtain for each firm a relationship between the value of the investment goods that it would plan to buy and the interest rate, and by adding these together, we will get a relationship between planned investment expenditure in each period by all firms and the rate of interest.

Our definition of the 'long-run' equilibrium rate of interest requires that this relationship between planned investment expenditure and the rate of interest and the economy's saving supply schedule must remain stable from period to period.

The figures in Diagram 88 portray an initial position in which the market rate of interest is at its 'long-run' equilibrium level. Figure (*a*) shows the speculative demand for money and the part of the total quantity of money that is available to meet it; Figure (*b*) shows the saving and investment schedules. At the rate of interest Oi, the part of their savings that the public wish to hold in the form of money is equal to the quantity of money that is available for acting as a store of value, and planned saving is equal to planned investment expenditure. Let us now suppose that at the beginning of period 1, this equilibrium is upset by a permanent rise in the speculative demand for money to D_1D_1, so that the market rate of interest rises to Oi_1. We shall suppose also that during the ensuing periods there is no change in (*a*) the tastes and preferences for present and future goods; (*b*) the prices of consumption goods and services; (*c*) the distribution of income; (*d*) the prices of productive services and durable goods; (*e*) the techniques of production and firms' awareness of them; (*f*) the quantity of money, and that the planned investment expenditures are independent of the level of the economy's income. At the new market rate of interest Oi_1, that



(a) (b)

DIAGRAM 88

rules at the beginning of period 1, planned saving for that period will exceed planned investment spending by *ab*. On our definition of a period, the income of the economy will fall by *ab* during period 1. This fall in income will mean that fewer units of money are required for transactions purposes, so that by the end of period 1 the number of units available to meet the speculative demand will have risen — from OM_1 to OM_2. If we assume that changes in the supply of money that is available for speculative purposes during any period only affect the market rate of interest at the beginning of the next period, then at the beginning of period 2 the rate of interest will be Oi_2.

As a consequence of the fall in the economy's income during period 1, the saving supply schedule for period 2 will be to the left of its initial position at S_{y_1}, for we assumed that saving is related to the previous period's income. At the market rate of interest of Oi_2, planned saving will exceed planned investment expenditure by *cd*, and during period 2 the economy's income will fall by this amount; this fall in income will reduce the demand for money as a medium of exchange, so that by the end of period 2 the quantity of money that can act as a store of value will have risen to OM_3. For period 3, the market rate of interest will be at Oi_3, and the saving supply schedule at S_{y_2}. During period 3, planned saving will exceed planned investment spending by *ef*; this will cause a further rise in the amount of money available to meet the speculative demand, and so a further fall in the market rate of interest. It can be seen from Diagram 88, that the reductions in income become smaller and smaller with each ensuing period, so that the reductions in the market rate of interest become smaller and smaller also. Eventually, the rate of interest will reach some new 'long-run' equilibrium level at Oi_n, at which planned saving (with the saving supply schedule $S_{y_{n-1}}$) will be equal to the planned investment expenditure.

In this analysis of the 'long-run' behaviour of the interest rate in response to some initial change, we have assumed that variation in the economy's real income is the sole equilibrator. Our analysis can be easily modified to allow for changes in some of the *cetera* that we have assumed to remain *paria*. Thus, if we assume that the investment schedule is not independent of income and posit some functional relationship between it and real

income, we may trace another path of adjustment of the rate of interest to a different 'long-run' equilibrium level. We may make the investment schedule simply dependent on the previous period's income as we did with the saving schedule; if we do so, the 'long-run' equilibrium level of the rate of interest will be lower than in our example. We may assume that the investment schedule depends on the rate of change of income in the recent past — that is, that investment expenditure in period t is a function not only of the interest rate but also of $Y_{t-1} - Y_{t-2}$; if we do so, we shall find that the interest rate will fluctuate over time, either converging towards, or diverging from, some 'long-run' equilibrium level.* In Diagram 88, we have assumed that all prices are constant (because all supply curves are perfectly elastic) so that changes in money incomes represent changes of the same proportion in real income; an alternative analysis of the long-run behaviour of the interest rate might be based on the assumption that the real income of the economy is stable, so that changes in money income represent changes only in prices. To do this, we must posit functional relationships (at the least) between money income and planned saving and investment and the demand for money as a medium of exchange. There will, therefore, be as many 'long-run' equilibrium rates of interest as there are 'long-run' equilibrators;† our aim is neither to catalogue them nor to choose between them, but merely to indicate one way in which the 'long-run' adjustments, with any given equilibrator(s), may be analysed.

With the aid of an analysis of the same kind as that illustrated in Diagram 88, we may offer a first approximation to an interpretation of the role that the productivity of investment goods and the tastes and preferences of savers play in determining the interest rate. A change in the former will shift the investment schedule, *ceteris paribus*; a change in the latter will shift the saving supply schedule at each level of real income, *ceteris paribus*. Such changes will affect the rate of interest in the model portrayed in Diagram 88 through changes in the economy's

* See P. A. Samuelson, 'Interactions between the Multiplier Analysis and the Principle of Acceleration', Chapter 12 in *Readings in Business Cycle Theory*, Philadelphia, The Blakiston Company, 1944.

† In Diagram 88, the main equilibrator is real income. An alternative equilibrator might be money incomes and prices. The equilibrating process may be assisted by changes in real or money investment.

real income. The detailed argument is left to the reader, for its form is similar to that described on pages 235–6 above.

Thus far in this section, we have assumed a world in which there are only two assets, namely, money and homogeneous, perpetual bonds. We shall briefly indicate how our analysis may be formally extended to a world in which there are n assets, A_1, A_2, A_3 ... A_n. The relationship between the prices of these will be determined by the demand for, and the supply of, each of them. We shall suppose that at any point in time the quantity of each asset is given, and that over rather short periods of time the amount by which the stock of any one of these assets can be augmented or depleted is negligible: the supply of each asset will then be perfectly inelastic. Given the stock of each asset, the demand for any asset, A_n, will depend on the public's tastes and preferences for it as compared with each of the others, and on the current market price of A_1, A_2 ... A_{n-1}. The demand curve for each asset will generally be relatively elastic, for it may be substituted for other assets, and others may be substituted for it, in response to changes in relative asset prices. In an equilibrium position, the relationship between the prices of the different assets will be such that the public, taken as a whole, will just be willing to hold the existing stock of each asset. If the equilibrium is upset, through a change in the public's preferences for some assets as compared with others, the demand curve for each asset will move to a new position as a consequence, and there will ensue a process of adjustment during which there will be further shifts in the demand curves in response to changes in relative asset prices, until a new equilibrium position is reached. Thus, if assets A_1, A_2, ... A_{12} are riskless bonds of progressively longer currencies, ranging from a three months' bill to an irredeemable bond, and if the public as a whole expects the general level of bond prices to be higher in future than they had previously thought, then the demand curve for each of these will rise, with that for A_{12} rising most and that for A_1 rising least, and the demand curves for money and other assets will tend to fall. These initial changes in the demands will alter relative asset prices and so lead to further shifts in the demands, and these will continue until, in the light of these new expectations about the future bond prices, the public are just willing to hold the given stock of each asset. In such a world, there will be no *the*

rate of interest: rather, there will be as many rates of return as there are assets. The rate of interest that any individual firm, *X*, must pay for a loan of money will depend, *inter alia*, on how potential lenders feel about *X*'s capacity to pay the interest and repay the principal, and on the period for which the loan is required. These will be reflected in the tastes and preferences of the public for the bond (asset) that *X*, the borrower, is selling. The price that *X* will get for his bond gives us the rate of interest that he must pay, and the price he can get will be the market price of those existing bonds that are in all respects identical with that which he is offering for sale.

In this way, we may explain the price that any firm *X* must pay for a loan of money for a given period of time. If we define interest as the price that is paid solely for the use of money, then the price that *X* pays will consist of more than interest, for those who sell the use of money to *X* are selling also their willingness to bear the risks of *X*'s default. We will get a rough notion of the part of the price that *X* pays that may be called 'pure' interest from the price that a riskless borrower (like a central government) pays for a loan of the same size for the same period of time. Our prime purpose in this chapter, however, is to explain the determination of the relative prices of the things that firms buy. In this pursuit, there is no need to break down the price of any productive service into such notional components as 'pure interest', 'rent' and 'wages'.

CONCLUSION

In this chapter, by combining the analyses of Chapters 5 and 6, we have attempted to explain why the price of a productive service might rise or fall as compared with the prices of products and of other productive services. In explaining the price of a particular kind of labour-service, for example, we have tacitly assumed that changes in the demand for, and in the supply of, it do not cause changes in other prices. Thus, if the demand for carpenters' services rises because there is a rise in the demand for the products they help to make, we have assumed that as a result of this rise in the demand for certain products the demand for no other product falls significantly so that there is no appreciable change in the demand for other productive services: on these assumptions, the prices of other products and factors re-

main unaltered, so that the demand for carpenters' services rises initially to its final level. An analysis of this kind that neglects causative or consequential changes in other markets when explaining changes in one particular market is called a partial or particular analysis. In explaining the behaviour of interest rates, however, such an analysis is uncomfortably constrictive: if at the beginning of some period, there is a permanent revision in wealth-holders' expectations about future interest rates so that the current rates fall, then in that period the demands for durable goods will rise, causing some increases in their prices. These changes will shift the supply curves of all products for whose production these durable goods are required, and so lead to changes in their prices; and so on. While a rise in market rates of interest is, when it occurs, a rise in them as compared with all other prices, there will be consequential changes in other prices, and these will react back on interest rates. The analysis of such consequential changes requires a general, rather than a particular, analysis, and this is the subject of the chapter which follows.

CHAPTER 8

The Determination of Relative Prices

We have now described the purchase and sales plans of individual households and firms, and we have shown why and how each of these plans will be revised if any datum on which it is based should alter. So far, however, we have explored only some of the consequences of a change in a planning datum: indeed, we have generally concentrated on its effects on the relative price of a particular product or productive service. In this brief chapter, we shall first, distinguish between partial and general analysis; second, attempt to describe the probable effects of an autonomous revision in the plans of some or all households or firms upon the general structure of relative prices, and third, discuss the usefulness of general analysis. In the next chapter, we shall examine the usefulness and realism of the model economy that we have analysed in this and in the previous seven chapters — namely, an economy wherein each price is both a datum for each buyer and seller and a resultant of the decisions of all buyers and sellers.

GENERAL AND PARTIAL ANALYSIS

In exploring the general consequences of an economic event, we shall be extending the analysis of the previous seven chapters; it may, therefore, be helpful to begin with a brief recapitulation to illustrate its main characteristics. It will be recalled that the data for the individual household as buyer, for example, are its tastes and preferences, the prices of all the goods that it might want to buy and its planned consumption expenditure; in making a purchase plan, the household is choosing the quantity of each good and service that promises the fullest satisfaction of its desires. In Chapter 1, we described how the planned purchases of any good X would be revised in response to changes in its own price, and we showed how this relationship (which we called the household's demand for X) would alter as a result of

changes in the household's tastes and preferences or in the price of any good other than X. The data for the individual firm as seller are the production possibilities that are open to it, the prices of its productive services and its products, and its objective; in making a sales plan, the firm is deciding what quantity of its products to sell in each period to earn the maximum net revenue. We have described how the planned sales of any good X would be revised in response to changes in its own price, and we have shown how this relationship (which we called the firm's supply of X) would respond to changes in the firm's production possibilities, in the price of any productive service that the firm uses or might use, or in the price of any product other than X. In Chapter 4, we showed how the demands of all buyers and the supplies of all sellers of X determined its relative price, *ceteris paribus*, and we described how the relative price of X would alter if there were any change in the tastes of the buyers, in the production possibilities open to the sellers, or in the price of any productive service or other product. In that way, we saw that the relative price of X, which is a datum for each individual household and firm in our model, is determined by all the households and firms that buy and sell X: that is, that which is a constant for each is a variable for all.

An analysis of this kind is called a partial or particular analysis. It is partial for at least two reasons. First, it confines itself to only some aspects of the economic behaviour of households and firms: it may, for example, limit itself to the individual household or firm as buyer or seller, or to all buyers and sellers of a particular commodity or productive service. Second, it is partial in that it explores only some of the economic consequences of the event that it describes: thus, in describing what might happen when households and firms implement their plans to buy and sell some good X, it concentrates on the behaviour of the relative price of X.

This kind of partial analysis has at least three advantages. First, it provides us with a broad classification of the causes of a change in the behaviour of any individual household or firm or of a change in the price of any individual commodity or service. Second, it helps us to predict the more immediate consequences (and these will generally be the most important consequences) of any general revision in the plans of those who buy

and sell a particular good or productive service, or of any interference by a political authority in the market for it. Third, by concentrating on one or a small number of economic subjects, we reduce the number of variables whose value must be explained and so simplify our analysis. Thus, when describing the purchase plan of the individual household, the dependent variable — that is, the quantity whose magnitude we are seeking to explain — is the quantity of each good, and the independent variables or data are tastes and preferences, consumption expenditure and prices; when describing revisions in its plan, we can assume a simple causal relationship between changes in the latter and changes in the former. Similarly, when exploring what happens as all buyers and sellers of commodity X implement their plans to buy and sell it, the dependent variables are the price per unit of X and the quantity of it that will be bought and sold per period, and the independent variables are the consumption and production possibilities and all other prices. Here, however, we are approaching, if not transgressing, the limits of partial analysis, for a change in the price of X that is caused by a change in one of the independent variables may exert a perceptible influence on this or on another 'independent' variable: thus, a rise in the demand for X may be associated with a fall in the demand for one or more other commodities, so that as the price of X rises that of other goods may fall, and this change in the pattern of relative prices will cause revisions in individual sales and purchase plans leading to further shifts in market demand and/or supply schedules until full adjustment has been made to the new conditions. When this happens and when it is significant, our distinction between 'independent' and 'dependent' variables — that is, between those things which determine and those which are determined — and the simple causal sequences which go with it, become unsatisfactory, for instead of simple dependence we have interdependence.

This interdependence becomes the more important the larger is the group of economic subjects in whose behaviour we are interested. For the individual household or firm, it may be insignificant; for all the firms in an industry and all the households that buy their product, it may be perceptible. If we take all those who make economic decisions in an economy, then the interdependence and mutual determination are paramount.

The nature and consequences of the inter-relation between economic quantities can best be seen by listing the general effects of some economic event, and we shall do this by using the analysis of the preceding chapters.

THE GENERAL CONSEQUENCES OF AN ECONOMIC EVENT

Let us suppose that the whole economy is initially in 'general' equilibrium: that is, that at the going prices the planned sales of each commodity and productive service are equal to the planned purchases. Let us now suppose that, for some extra-economic reason, the preferences of households for commodity X become stronger. We may obtain a clearer picture of the nature and importance of economic interdependence by tracing the consequences of this event through the whole economy. The order in which these consequences are listed below is that which follows most easily from the application of our previous analysis; it is not the order in which they occur in calendar time, and they are not listed in the order of their relative importance.

(*a*) The demand curve for X and the short-run demand curves for the productive services that assist in its production will all shift rightwards. The price of X, and the price of each variable productive service that the firms that produce X require, will all tend to rise. These increases will tend to be the greater the less is the elasticity of the short-run supply curve of each relevant productive service and the more rapidly does the rate of increase in the physical output of X decline as more and more of the variable services are combined with the fixed factors at the disposal of the firm.

(*b*) While these tendencies in (*a*) are operating, other changes will be occurring. If we assume that the initiating change in preferences is not associated with any alterations in the total consumption expenditure of households, then when the demand for X rises, the planned expenditure on all other goods taken together must decline. If the expenditure on buying X in the new conditions of demand for it is still only a very small proportion of total spending, and if the increased spending on X is equally at the expense of all other goods, then the demand curve for none of them would shift significantly, so that their prices would remain unaltered and the partial analysis under (*a*) above would suffice. If neither of these conditions is fulfilled,

then when the demand for X rises, the demands for at least some other goods must fall, and we would expect the demand for goods that are substitutes for X to fall, and that for goods that are complementary to X to rise. These changes in the demands for other products will cause changes in the same direction in the short-run demands for the variable productive services needed to produce them. The prices of substitutes and of the factors that assist in producing them will tend to fall; the prices of complements and of the factors needed to make them may tend to rise.

(*c*) The changes listed under (*a*) and (*b*) above will affect one another. The tendency for the prices of substitutes for X to fall, and the prices of complements to X to rise, will tend to inhibit the increase in the demand for X, for the initial rise in the demand for X was assumed to be solely the result of the change in preferences. The tendency for the price of X to rise will lead to some increase in the demand for substitutes for X and some fall in the demand for complements to X, though in neither case will these changes offset the initial changes in demand resulting from the change in tastes. If the productive services used in the X and X-complement industries are quite different from those used in the X-substitute industries, then during the short-run there will be no inter-relations between their supplies. If the factors are similar, then as a consequence of the fall in the prices of the X-substitute factors, the supply curve of the same or similar factors to the X and X-complement industries will shift rightwards, and as a result of the rise in the prices of the factors in these latter industries, their supply to the former will decline. The net effect will be that costs of production will rise by less in the expanding industries and fall by less in the contracting industries.

(*d*) Thus far, we have confined our attention to the short-run. In the long-run, as a consequence of the changing pattern of relative product and factor prices entrepreneurs may decide to produce different products or to produce their existing products by other methods. And these decisions, when implemented, may lead to further changes in prices. The changing pattern of relative wage-rates, for example, may induce labourers to discard their old skills and acquire new ones, and this will affect the supplies of the different kinds of labour-service and so lead to further changes in prices. Over the long-run, therefore, the re-

percussions of the initial change in households' preferences will be much wider, for an entrepreneur who decides to begin producing X or a labourer who acquires a new skill that enables him to assist in its production may have been employed previously in almost any other industry in the economy.

(*e*) The change in the relative prices of productive services may cause changes in the distribution of income between persons, and since tastes and preferences vary from one individual to another, this will cause changes in the pattern of demand for consumption goods and services, and so in the demands for productive services. The changing personal distribution of income may lead to changes in planned saving, and this may tend to change the level of the economy's income and in that way exert some influence on the pattern of interest rates. And changes in the level and pattern of interest rates will affect the demands for productive services.

The prices of all products are bound together by the possibilities of substitution; the prices of productive services are related to one another through substitution and by the ability of some of them to change their form — thus, a carpenter may become a bricklayer, the accumulated amortisation on a weaving machine may be used to buy a spinning machine; and the prices of products are linked to the prices of productive services by the possibilities of firms transforming the latter into the former *via* production. In the long-run, there is full scope for all substitutions and transformations, which appear to be profitable at the existing pattern of relative prices, to be effected, and as they are effected to induce further changes in relative prices. It is probable that when all these long-run consequences of the initial autonomous rise in the demand for X have worked themselves out, the prices of X and of the productive services used in making it, will be relatively higher than they were initially, though not so high as they were in the short-run.

The preceding paragraphs, though sketchy and cursory, must suffice to illustrate the ramifications of some initial change in households' preferences. When a 'general' equilibrium is disturbed by some economic event, there will ensue a process of adjustment and readjustment during which each price affects, and is in turn affected by, each other price. The changing pattern of prices is in part the cause, and in part the consequence,

of revisions in the purchase and sales plans of individual households and firms. During the adjustment towards a new equilibrium position, each household and firm lays its purchase and sales plans on the basis of the existing pattern of relative prices;* as these plans are implemented, the pattern of relative prices alters, leading to further revisions in plans that cause further changes in relative prices. Ultimately, a new equilibrium will emerge, in which the planned sales of each commodity and productive service will again be equal to the planned purchases. In the new 'general' equilibrium, full adjustment will have been made to the new conditions. In our example, there will be a new pattern of relative product and factor prices; the proportions in which the different productive services are being used by firms will have altered; the distribution of skills amongst the different members of the labour-force and the composition of the economy's stock of durable goods and equipment may both have changed, and there will have been changes in the composition of the flow of goods and services that is being produced in each period and in the manner in which this flow is being distributed between households.

USES OF GENERAL ANALYSIS

The general analysis of the previous section has at least four main uses: first, it helps us to predict the general consequences of an economic event; second, it typifies the nature of economic interdependence;, third, it illustrates the general functions of prices, and fourth, it helps us to identify the largely non-economic determinants of the pattern of economic activity. Since the first two of these uses have already been illustrated by the example that we worked through in the previous section, we may deal with them briefly. With the aid of general analysis, the ramifications of any autonomous economic event — that is, an event that is not itself caused by a change in price(s) — may be unravelled; the analysis helps us to list the economic quantities that will probably be affected and to predict (with some confidence) the direction in which their magnitudes will alter. In exploring the general consequences of any autonomous event, we are not merely listing the steps in a simple causal sequence, for

* Or of the pattern of relative prices that is expected to obtain when these plans are being implemented.

while event A may cause event B, a part of the consequence of B may be a change in the intensity with which A operates. Thus, an initial rise in the demand for commodity X may cause a rise in its price and some reductions in the prices of other commodities, and these in their turn may lead to some contraction in the demand for X; and in the longer run, the demand for X may be further affected if there is any significant change in the prices of productive services or in the distribution of income. While our general analysis helps us to list the consequences of some initial event in an interdependent economy, and to arrange them roughly in the order in which they are likely to appear, we must use our own judgement in deciding which consequences are significant and which trivial. It is probable that for most economic events, the more important consequences can be predicted with the help of partial analysis.

The third use of general analysis is in illustrating the function of prices. In our model economy, wherein each commodity and each productive service has a price that varies in response to changes in the demand for and in the supply of it, individual economic decisions are integrated and made consistent with one another through changes in relative prices. The manner in which this integration is effected has been illustrated repeatedly in Chapters 4 and 7 and in the previous section of this chapter. Thus, if a group of households plan to buy more of some commodity X, then planned purchases of X will be inconsistent with the planned sales of X by the firms that produce it; in the market period, these plans are rendered consistent with one another by a rise in the relative price of X. When this happens, however, the sales and purchase plans of the firms that produce and sell X will become inconsistent with each other and with the firms' objectives. In the ensuing short-period, the firms will revise their purchase plans within the limits set by the fixed factors at their disposal; as they do so, however, the total planned purchases of these variable productive services will become inconsistent (at their existing prices) with the sales plans of the households that supply them. The purchase and sales plans of productive services will, in their turn, be made to coincide with one another through changes in the prices of productive services. And so on: the changes in relative prices and the revisions in the purchase and sales plans of households and firms (of which the former are

both the cause and consequence) will continue until the planned purchases and sales of each commodity and productive service are equal to one another.

As changes in relative prices render individual economic plans compatible with one another, three broad decisions are made for the economy as a whole: first, what commodities to produce, and in what quantities; second, how (that is, with what quantities of what productive services) to produce them, and third, what households shall have these goods once they are produced. The first two of these broad decisions are the resultant of the decisions of each individual firm to implement the sales and purchase plans that promise it the maximum net revenue. In choosing these plans, each firm is guided by the relative prices of products and by the relationship between these and the prices of productive services. The relative prices of the products that a firm might produce reflect directly the relative intensity of the demands by households for them, and the relative prices of productive services reflect the relative intensity of firms' demands for them, and these, as we know, are largely determined by the demands of households for the products that the productive services may assist in producing. Since the tastes and preferences of households play a decisive role in determining what goods shall be produced and in what quantities, the patterns of production and resource-use that emerge will correspond to those that households prefer. The third broad decision — the distribution of the total product amongst households is made as households make their tastes and preferences effective. Households help to mould the pattern of production in the economy, more or less in accordance with their desires, by the manner in which they spend their incomes. Since the influence that any particular household exerts depends on the size of its income as compared with that of other households, its relative share of the economy's total product will be the same as its relative power to determine the constitution of the total product.

Our model economy, wherein all prices are both data for each individual buyer and seller and variables for all buyers and sellers taken together, typifies one hypothesis about the causation of human events, namely, that it is 'only by taking an infinitesimally small unit of observation (the differential of history, that is, the individual tendencies of men) and attaining to the

art of integrating them (that is, finding the sum of these infinitesimals) . . . [that we can] . . . hope to arrive at the laws of history'.* The differentials of economics are the decisions of individual buyers and sellers; the fact that each thing that any buyer might want to buy, or any seller to sell, has a price that responds to changes in the demand for, or in the supply of, it helps to integrate the myriad of individual decisions and so determine the course of economic events. That we can explain the causes of economic events, and predict (and with some confidence) the course of future economic events, is due more to the existence of certain economic institutions than to the inspiration of economists, for what economists can do, that historians and political (and other social) scientists cannot do, is explained largely by the existence of prices. The practice of attaching a money value to each commodity and productive service spontaneously developed in societies in which there was private ownership of the factors of production and of the services they rendered and of the goods they produced, and in which each buyer and seller had a tolerably wide range of choice in deciding what, where and when he would buy and sell. We may, of course, still have prices where both private property and freedom of economic choice are absent, but these prices are not the medium through which individual economic decisions are integrated, but rather a part of the mechanism by which broad economic decisions, that have been arrived at by some political (or other) process, are implemented. In such societies, the causes of economic events and of their future course must alike elude the economist, and for the same reason — namely, the absence of any overt and measurable mechanism by which individual decisions are integrated with one another — as the causes and consequences of political and social events may elude political and social scientists in capitalistic societies.

In the fourth place, our model economy helps us to understand the ultimate determinants of the pattern of economic activity. Since each household helps to determine what goods shall be produced by the manner in which it spends its income, the pattern of production in the economy will reflect the distribution of tastes and of income between households. We accept

the pattern of tastes as a datum, for the wants of households and the kinds of goods that they believe will satisfy them are largely determined by the physiology of human beings and the society in which they live. The income of any household depends on the quantity of productive services that it owns and on the price at which it is currently selling them: for the economy as a whole, the pattern of income distribution will depend, then, on the distribution of the ownership of productive services (which we shall call, for brevity's sake, 'property') between households and on the current prices of productive services, which are in part determined by the prices of the products they help to produce, and these, in their turn, reflect the pattern of households' tastes and the extent to which the households can make them effective in shaping production. For short periods of time, the extent to which the firms in an economy can satisfy the pattern of household demand will ultimately depend on three things: first, the number of workers, the quantity of land and the stock of physical and potentially productive equipment available to the economy as a whole; second, the versatility of each unit of labour, land and equipment — that is, the different products in whose production it might assist, and third, on the laws of the physical sciences and the state of the technical arts which together delimit the range of combinations of productive services by which any particular product might be produced. Over rather long periods of time, some of these may be as much consequences as causes of the pattern of prices: thus, there may be some relationship between the size of the population (and therefore of the potential labour-force) and money wage-rates, or between it and the relationship between money wage-rates and product prices (that is, 'real' wages); the composition of the stock of equipment may respond to changes in the relationship between wage-rates and the prices of the services of durable goods, and its size to changes in the relationship between interest rates and other prices; and it is possible that the expansion of technical knowledge and certainly the rate at which it is applied, will both depend on changes in the pattern of relative prices.

In the short-run, therefore, the pattern of demand in an economy in which there is a freely operating price mechanism reflects the pattern of property ownership and the pattern of tastes, and the pattern of supply depends on the quantity and

the quality of the productive services available to the economy and on the state of the technical arts. The supply of productive factors and the extant technology together delimit the patterns of production that an economy might have; prices are the medium through which households decide what particular pattern of production it shall have. If the pattern of production that is chosen is deemed undesirable, then our disapprobation must be directed either to the pattern of tastes or to the distribution of income or to both of these. The pattern of production may be brought slowly into accord with what is thought desirable by altering tastes through education and example, and by modifying the distribution of property between households by inheritance taxes or the distribution of income by income taxes. If it is desired to reduce the production of some commodities quickly, this may be done by imposing specific taxes on them to raise their relative prices. The cause of the unacceptable pattern of production cannot lie, in our model, in the price mechanism itself, for it is merely the medium through which the effective desires of households are transmitted to firms: to seek a remedy by eliminating the price mechanism would be like destroying a mirror because it gave a reasonably true image.

It might happen, however, that the desired changes in the pattern of production are so drastic that the changes in tastes or income-distribution needed to achieve them are either psychologically or politically impossible, or productive of consequences that would be deemed undesirable. Thus, a radical redistribution of wealth, designed to make the patterns of consumption and production more desirable, may reduce the incentive to indulge in economic activity in order to acquire wealth and so ultimately reduce both total production and total consumption. Or again, the current distribution of production between goods for current consumption and investment goods (that is, goods to make goods to augment the consumption flow in the future) may promise a rate of increase in total production or consumption that is unacceptable, and it may be that the disposition of households' incomes between spending and saving that is appropriate to the desired rate of expansion cannot be achieved when there is a freely operating price mechanism and democratic political institutions. In these circumstances, it may superficially seem that there is a conflict between economic ends and the means to

their achievement. In effect, however, there is a conflict between ends, for the institutions on which a freely operating price mechanism rests are 'ends' in themselves: the reasonably wide freedoms to choose one's job and what goods to buy and sell are as important as the freedom of speech or the freedom to choose one's rulers. Indeed, the former have tended to precede the latter in the course of history, and the relationship between them may be causal rather than accidental. In passing judgement on any economy, therefore, we must not pass judgement solely on its results, for the methods by which these are reached may be equally relevant.

MATHEMATICAL SUMMARY

In this final section, we shall summarise algebraically our conclusions thus far. We shall use 'p' to represent price, 'q' to represent quantity, and the subscripts 1, 2, 3, ... n to designate particular *commodities* — that is, things which households buy and firms sell. We shall use 'f' to denote price, 'x' to denote quantity, and the subscripts 1, 2, 3 ... m to designate particular *productive services* — that is, things which households sell and firms buy. We shall use 't' to denote the rate at which productive services may be transformed into commodities: thus, if we are interested in commodity i, we shall denote the quantity of productive service j being used to produce each unit of it by t_{ij}. Our main conclusions can be translated into this notation as follows:

(i) Demand for commodities: Given the tastes, preferences and aims of the households in the economy, the quantity of any commodity i that will be demanded will depend on its own price and on the price of each other commodity and on households' incomes which in turn depend on the prices at which they are selling their productive services. This proposition may be written as follows:

$$q_i = F_i(p_1, p_2, p_3, \ldots, p_n, f_1, f_2, f_3, \ldots f_m) \qquad (1)$$

where F_i denotes the nature of the dependence of the quantity purchased on the prices of commodities and productive services.

(ii) Cost of production of commodities: Given the market morphologies,* the state of the technical arts and the objectives of firms, the price at which any commodity i will be sold will depend on the costs of producing it, which in turn depend on the quantities of the various productive services used and on the prices that have to be paid for them. That is:

$$p_i = \phi_i (t_{i1} \cdot f_1, t_{i2} \cdot f_2, t_{i3} \cdot f_3, \ldots t_{im} \cdot f_m) \qquad (2)$$

where $t_{i1} \cdot f_1$, for example, represents the money spent on buying the quantity of productive service 1 which must be used to produce one unit of commodity i, and where ϕ_i denotes the nature of the dependence of price on costs of production. If each of the n commodities is being produced under conditions of pure and perfect competition,† then the price of each commodity will be equal to the average total costs of producing it — that is:

$$p_i = t_{i1} \cdot f_1 + t_{i2} \cdot f_2 + \ldots + t_{im} \cdot f_m \qquad (2a)$$

(iii) Supply of productive services: Given the attitudes, preferences and objectives of each owner of productive services, the kind and quantity of any service j that he will sell will depend on the relationship between the prices of j and of all other productive services and on the relationship between these and the prices of commodities. That is:

$$x_j = F_j (f_1, f_2, f_3, \ldots, f_m, p_1, p_2, p_3, \ldots p_n) \qquad (3)$$

where F_j denotes the precise nature of the relationship between planned sales and prices.

(iv) Demand for productive services: Given the state of the technical arts and the aims of firms, the quantity which firms will use of any productive service j in producing each unit of commodity i will depend on the relationship between the prices of j and of all other services — that is:

$$t_{ij} = \pi_j (f_1, f_2, f_3, \ldots, f_m) \qquad (4)$$

where π_j denotes the nature of the dependence of the quantity of j demanded on the relative prices of productive services.

* See *infra*, Chapter 9.
† *Ibid.*

(v) Finally, if there are no unemployed productive services, then the quantity of any service j that is being supplied must be the same as the quantities of it that firms are buying — that is:

$$x_j = t_{1j} \cdot q_1 + t_{2j} \cdot q_2 + t_{3j} \cdot q_3 + \dots + t_{nj} \cdot q_n \tag{5}$$

where $t_{1j} \cdot q_1$ means the quantity of j used in producing commodity 1, and $t_{2j} \cdot q_2$ the quantity of j used in making commodity 2, and so on.

These five equations are a convenient shorthand in which we can summarise the main relationships described in this and in the previous seven chapters. If we substitute for i the values 1, 2, 3, ... n, and for j the values 1, 2, 3, ... m, we get five groups of equations which together describe the economy. If the prices of all commodities and productive services are measured in terms of commodity 1, then 1 will be the *numéraire* and its price will be unity. The demand equation for commodity 1 must then be modified: if we suppose that each household spends the whole of its income, then the demand for the *numéraire* will be equal to the total income of households less their expenditure on commodities 2, 3, ... n. The first group of equations will then be as follows:

$$\begin{aligned} q_1 &= f_1 \cdot x_1 + f_2 \cdot x_2 + \dots + f_m \cdot x_m - p_2 \cdot q_2 - p_3 \cdot q_3 - \dots \\ &\qquad - p_n \cdot q_n \\ q_2 &= F_2(p_2, p_3, \dots p_n, f_1, f_2, \dots, f_m) \\ &\quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ q_n &= F_n(p_2, p_3, \dots, p_n, f_1, f_2, \dots, f_m) \end{aligned} \tag{I}$$

If there is universal pure and perfect competition, the cost of production group of equations will be:

$$\begin{aligned} p_1 = 1 &= t_{11} \cdot f_1 + t_{12} \cdot f_2 + \dots + t_{1m} \cdot f_m \\ p_2 &= t_{21} \cdot f_1 + t_{22} \cdot f_2 + \dots + t_{2m} \cdot f_m \\ &\quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ p_n &= t_{n1} \cdot f_1 + t_{n2} \cdot f_2 + \dots + t_{nm} \cdot f_m \end{aligned} \tag{II}$$

We shall have m equations in group III describing the supplies of productive services, and mn equations in group

IV describing the demands for them in terms of the technical coefficients. Finally, we shall have the following m equations in group V:

$$
\begin{aligned}
x_1 &= t_{11} \cdot q_1 + t_{21} \cdot q_2 + \ldots + t_{n1} \cdot q_n \\
x_2 &= t_{12} \cdot q_1 + t_{22} \cdot q_2 + \ldots + t_{n2} \cdot q_n \\
&\;\;\vdots \\
x_m &= t_{1m} \cdot q_1 + t_{2m} \cdot q_2 + \ldots + t_{nm} \cdot q_n \qquad \text{V}
\end{aligned}
$$

One of the equations in these five groups — namely, the first equation in group I — is not independent, for it can be derived from the equations in groups II and V. If we multiply the successive equations in II by $q_1, q_2, q_3, \ldots, q_n$ respectively, and the equations in group V by $f_1, f_2, f_3, \ldots, f_m$ respectively, and add the equations in each group together, we shall find that the sum of the right-hand sides of each set of equations is the same. It follows that their left-hand sides must be identical — that is:

$$
\begin{aligned}
q_1 + p_2 \cdot q_2 + p_3 \cdot q_3 + \ldots + p_n \cdot q_n = f_1 \cdot x_1 + f_2 \cdot x_2 + \ldots \\
+ f_m \cdot x_m,
\end{aligned}
$$

$$
\begin{aligned}
\text{or:}\quad q_1 = f_1 \cdot x_1 + f_2 \cdot x_2 + \ldots + f_m \cdot x_m - p_2 \cdot q_2 - p_3 \cdot q_3 - \ldots \\
- p_n \cdot q_n,
\end{aligned}
$$

which is the first equation in group I.

In groups I to V, we have then $n + n + m + mn + m - 1$, or $2n + 2m + mn - 1$, independent equations. From these equations we want values for the $n - 1$ prices* and n quantities of commodities, the m prices and m quantities of productive services, and the mn technical coefficients. The total number of unknowns is therefore $2n + 2m + mn - 1$. Under certain simplifying assumptions, this equality between the number of independent equations and the number of unknowns means that the solution of the equations will give us a single value for each price, quantity and technical coefficient.†

This mathematical summary of our main conclusion has at least two uses. First, it provides, as we have seen, a convenient and brief way of restating the crucial economic relationships. Second, it tests the consistency of our analysis and the determinacy of the solutions it proffers. In describing the general consequences of an economic event in the second section of this chapter, we tacitly assumed that, given tastes, techniques and objectives, there existed a set of equilibrium values for the prices and quantities of commodities and productive services and the techniques by which the latter might be transformed into the former — i.e. sets of values for the prices and technical coefficients such that in each period the planned purchases of each commodity and productive service were equal to the planned sales. There, we concentrated on listing the changes in prices and quantities, and in the techniques used, which would follow any change in tastes, techniques or objectives. By writing the crucial economic relationships as equations and describing our model economy in terms of a system of simultaneous equations, we can verify that there exists at least one set of equilibrium values for the unknowns in the model.

CHAPTER 9

Market Behaviour and Market Morphology

Throughout the previous eight chapters, we have assumed that prices are data for each person (whether private or corporate) who makes economic decisions. This is an assumption about the way in which each individual household or firm behaves in the markets in which it buys and sells: each buyer and each seller is assumed to be a 'price-taker' or 'quantity-adjuster' — that is, in responding to changes in prices, each household and firm revises its purchase and sales plans by adjusting the quantities of the things that it buys and sells. On the basis of this assumption, and of the assumptions that in the course of its economic activity each household seeks the fullest satisfaction of its desires and each firm the maximum net revenue, we have constructed a model economy. Our purpose in doing so was to answer the question: why do relative prices vary? Within the framework of the model, this question has been answered, for the model provides us with a broad yet meaningful classification of the causes of changes in relative prices and a framework within which we may predict the probable consequences of present economic events.

From this model, we expected two things, namely, logical consistency and usefulness. We attempted to fulfil the former expectation by the manner in which the model was elaborated. The test of the usefulness of the model is the empirical validity of the explanations and predictions that it offers. If we take any actual change in relative prices — for example, the fact that twenty-five years ago a 'family' motor-car was roughly as expensive as a semi-detached house whereas it is now only one-quarter or one-fifth as dear — we generally find that the changes in tastes, production possibilities, etc., that have accompanied the changing pattern of relative prices are those to which our model attributes a causal role. If we use the model to 'predict' the consequences for relative prices of some past economic event

— such as Cort's invention of rolling and puddling, Hargreave's invention of the 'spinning jenny', the imposition of customs duties on imported commodities, or the introduction of a guaranteed minimum price for wheat in the Irish Republic — we generally find that its 'predictions' are in accord, in so far as the direction of price change is concerned, with what has actually happened. When predictions have been hazarded about the future consequences of current events, the passage of time has generally confirmed them.

Our model is useful, then, because it enables us to make hypotheses about the causes of changes in relative prices and the consequences of economic events for the pattern of relative prices. These hypotheses are empirically valid, and in that sense we may say that they are useful or realistic. Since the model was merely the elaboration of the implications of certain assumptions — notably, that each firm and household was a 'price-taker' when buying and selling — these assumptions must be judged useful and realistic also. The usefulness and realism of the assumptions is thus derived from the usefulness and realism of the hypotheses that they yield about the behaviour of relative prices, and not vice versa. An assumption that we judge realistic on this test need not be realistic in the sense that it is descriptively accurate — that it is a photographic representation of an actual, existing economy. Indeed, the more realistic is it in the former sense, the less realistic will it be in the latter, for assumptions that yield useful hypotheses are usually heroic simplifications of the observed reality. From all aspects and circumstances of the reality, such assumptions abstract those elements that are important for the phenomena that are to be explained. The only test of whether or not the assumptions that are made have succeeded in isolating the most important elements is whether or not the hypotheses they yield are confirmed by events.*

This thesis — that our judgement of the relevance and realism of assumptions must depend on the realism of their implications — has been laboured, because its converse has played a decisive role in shaping the development of the theory of price

during the past thirty years. The assertion that assumptions must be accurate descriptions of economic reality if they are to yield hypotheses that are empirically valid is implicit in Joan Robinson's *The Economics of Imperfect Competition* (London, Macmillan), and in E. H. Chamberlin's *The Theory of Monopolistic Competition* (Harvard University Press) — two works that have since acted as 'pathfinders' for the intellectual flights of many economists. If the phenomenon that they sought to explain was the behaviour of relative prices, their explanations add little to those to be found in Alfred Marshall's *Principles of Economics*, of which the previous eight chapters are largely a translation into a different idiom. 'The new theories [of monopoly, monopolistic competition, oligopoly, etc., which we shall study in the chapters that follow] seem to make little difference to the laws of change as they are exhibited in the traditional (that is, Marshallian) analysis. . . . Whether an industry is monopolized, or duopolized, or polypolized, or operates under conditions of perfect competition, we shall expect a rise in demand to lead to a rise in output . . . and it is still likely that the rise in demand will be accompanied either by no change in price, or by a rise. . . . It is therefore hard to see that the new analysis does much to displace Marshallian methods. Marshall's assumptions are simpler, and if we are unable to tell which of two hypotheses is more appropriate, the simpler has the obvious claim to be chosen.'* These new theories, indeed, provide proof that assumptions should be judged by the realism of their implications, for they show that the implications (for the behaviour of relative prices) of assumptions that themselves are empirically valid, are no more realistic or useful.

While the attempts to verify the assumptions of the traditional analysis — or rather to discover the extent to which they diverged from the real world — yield no more useful hypotheses than those that were already available, there did result from them a more detailed, and probably more useful, classification

of actual market structures or market morphologies. Consequently, we now have a number of boxes in which we may file actual firms and industries. It is on this aspect of these new theories that we shall concentrate in the following chapters. In the remainder of this chapter, we shall list the criteria on which the classification is based, and we can most easily discover these by asking: what would a real market have to look like if each buyer and seller in it were to act as a 'price-taker' and 'quantity-adjuster'?

PURE COMPETITION

The price of any commodity X will be a datum for each household that buys it and for each firm that sells it if the following conditions are simultaneously fulfilled:

(i) The number of sellers (firms) of X must be so large that the amount that each seller offers for sale in each period constitutes so small a proportion of the total quantity being supplied by all sellers that he, acting alone, is powerless to affect the price by varying the amount that he offers for sale. This number is incapable of being expressed cardinally. We can only define it operationally: the number of sellers must be so large that any practicable variation in the planned sales of any seller will not shift the market supply curve by enough to cause a change in the price, in the given conditions of demand.

(ii) The number of buyers (households) of X must be so large that the planned purchases of any buyer at any price constitutes an insignificantly small proportion of the total planned purchases at that price. This assumption is necessary for the same reason as (i) above, namely, to eliminate any appreciable or significant interdependence between the decisions of different buyers.

(iii) The product X that is being bought and sold must be homogeneous. The product will be homogeneous if each buyer (seller) is indifferent as to which unit(s) of a seller's production (buyer's purchases) he buys (sells), and if each buyer (seller) is indifferent as between sellers (buyers). If buyers are to regard each unit of X as being a perfect substitute for each other unit of X, then it is not only necessary that the different units of X must be physically and chemically identical; in addition, the spatial

distribution of buyers and sellers within a geographical area must be such that no preference can arise for reasons of distance for the product of any seller, and the circumstances that surround the buying and selling of X must be identical for all transactions — for example, all sellers must be equally polite or equally rude.

(iv) Each buyer, acting independently, seeks the fullest satisfaction of his desires within the limits set by his income and wealth, and each seller of X, acting independently of other sellers, attempts to earn the maximum net revenue per period by producing and selling it. This assumption is probably necessary to ensure that, when conditions (i), (ii) and (iii) are fulfilled, each buyer and each seller in fact behaves as a price-taker: the desire to maximise satisfaction or net revenue impels those who operate in the market to acquire enough knowledge about its main characteristics to realise that the price of X lies beyond the control of each of them. This assumption is certainly necessary if we are concerned with the efficiency with which a market in which conditions (i) to (iii) are fulfilled transforms productive services into products and distributes the products amongst households — that is, if our interest lies in welfare economics rather than in positive economics, if we are trying to answer the question: how well does the price system work? and not simply the question: how does it work?

(v) The existence of a market in which there are large numbers of buyers and sellers of a homogeneous commodity, with each buyer bent on maximising his satisfaction and each seller his net revenue, does not by itself guarantee that each and every purchase and sale of the commodity will be transacted at the same price. From our observations of markets wherein no individual buyer or seller can appreciably affect the price and wherein all transactions are effected at the same price — and the markets for wheat, cotton, coffee, and some other agricultural products provide approximate examples in the absence of government (or other) controls and regulations — we see that assumptions (i) to (iii) are always fulfilled. While these assumptions are necessary, they are not sufficient, conditions for each transaction being effected at the same price, and the markets for old gold and used jam-jars provide (or have provided) rather crude and approximate examples of their insufficiency. In addition to assumptions (i) to (iii) above, we must make some assump-

tion about the amount of knowledge that each buyer and seller must possess to ensure that all units of the commodity are sold at the same price. And given this assumption and assumption (iv) above, we know that this price must also be an equilibrium price.

It is customary to assume that all buyers and sellers must have complete knowledge of all prices and all price offers if all transactions are to take place at the same price. If we accord this 'perfect' knowledge to each buyer and to each seller, however, we make it impossible for the market in which they operate *not* to be in equilibrium, and we remain in ignorance about the way in which the equilibrium comes to be established. What is wanted, rather, is an assumption that answers the question: how much knowledge about what things must each buyer and seller possess if their joint actions are to be successful in establishing an equilibrium? We shall attempt to answer this question by examining the short-run equilibrium described in Chapter 4. The objective facts that underlie an equilibrium in the market for a particular product X are the tastes of each buyer, the production possibilities open to each seller (which depend on the physical productivities of the productive services that he uses), the price of each variable productive service, and the price of each product other than X. These define the environment of the market for X, and any equilibrium in that market must be relative to that environment. The equilibrium that actually emerges will depend on the knowledge that each buyer and seller possesses of these facts, and there will be as many equilibria as there are degrees of knowledge. If the equilibrium is to reflect fully all these facts, then each buyer must be aware of (*a*) his own tastes; (*b*) product prices, and (*c*) the relative capacity of different products to satisfy his desires, and each seller must be aware of (*a*) the quantity of X that he will actually obtain from any possible combination of the relevant productive services, and (*b*) the price of each productive service. The equilibrium that we described in Chapter 4 was of this kind: it was an 'ideal type' chosen to simplify our analysis by giving us a unique and determinate equilibrium, and it was also an 'ideal' of welfare economics. If we wish this equilibrium to be established in our model, then we must assume that each buyer and each seller possesses full knowledge of each of the things listed above.

In any actual market where conditions (i) to (iv) above are fulfilled, however, households and firms may possess less than complete knowledge of the relevant data, and the degree of incompleteness of their knowledge will vary from one market to another. If we want some assumption to guide us when dealing with such markets, other than the assumption of 'complete' or 'perfect' knowledge, we may assume that each buyer and each seller possesses that knowledge of the relevant data that he is bound to acquire as he implements and revises his plans. Thus, a household may make its initial purchase plan on the basis of certain expectations about the capacity of certain products to satisfy its wants; as it actually buys and consumes these products it will gain some clearer notions about the satisfactions they provide, and its purchase plan for the subsequent period will be laid on the basis of these. Similarly, a firm will acquire a fuller knowledge of the actual physical productivities of the productive services it uses as it compares, and seeks to account for, the difference between the net revenue it expected to earn and that which it actually succeeded in earning. The knowledge that is acquired in this way will be the fuller, the more stable are the elements of the market environment. Even if the environment is absolutely stable, however, the knowledge that is so acquired need never be complete, for there may be certain relevant data that a household or firm never becomes aware of through putting its plans into effect. Thus, if a household is initially unaware of the existence of commodity *Y*, which would satisfy the same want as *X*, *Y* will not appear in its purchase plan and the implementation of that plan will not necessarily call *Y* to the household's attention. For the present, however, we shall assume that each buyer and seller has complete knowledge of his 'segment' of the market environment: this simplifies our analysis in that it gives us a unique equilibrium price for market *X* when conditions (i) to (iv) obtain.*

When the structure of the market for a commodity is accurately described by the five assumptions that we have listed above, we shall say that those who buy and sell in it are operating under conditions of *pure competition*, or that in that market

* On our assumption (v), see F. A. Hayek, 'Economics and Knowledge', *Economica*, IV, New Series, 1937, pages 33–54.

there exists a state of pure competition. When a state of pure competition exists in the market for X, each purchase and sale of X takes place at the equilibrium price and each household is buying in each period a quantity of X such that its satisfaction from its total consumption is maximised, and each firm is selling the quantity of X that promises it the maximum net revenue per period; alternatively, we may say that each household's purchases of X are such that the marginal rate of substitution between X and each other product is equal to the ratio of their prices, and each firm's sales of X in each period are such that (*a*) the marginal cost of production is equal to the equilibrium price, and (*b*) the marginal rate of technical substitution between any two of the productive services used to produce this quantity of X is equal to the reciprocal of the ratio of their prices.

In delineating the morphology of a purely competitive market, we have confined our attention to a market for a product. Pure competition may also exist in the market for a productive service — that is, in the market for something that firms buy and households (or other firms) sell. The assumptions that must be fulfilled to give us a purely competitive factor market are very similar to those listed above, and a statement of their precise contents is left to the reader. It should now be clear that in our explanation of the determination of product and factor prices in Chapters 4 and 7, we assumed that a state of pure competition existed in the product and factor markets respectively.

PERFECT COMPETITION

If pure competition exists in the market for some commodity X, then we know that at each moment of time each buyer and seller of X will be a price-taker, and the price of X will be such that the planned purchases of all buyers will be the same as the planned sales of all sellers. If the market environment alters, however — as a result, for example, of a general change in the intensity of the desires for X — the existence of pure competition tells us nothing about the relationship between the prices of X at successive moments of time: it tells us merely that at each instant the price is such that the market is cleared; it does not tell us anything about the path that will be traced by the price of X as time passes. We saw in Chapter 4 that the time-path of prices following some initial event depends primarily on the

adjustments that are effected in the firm's sales plans. We classified all the adjustments that might occur into two groups, namely, short-run and long-run adjustments. In the short-run, the only revisions of sales plans that are possible are those that can be made by the firms already producing the product, within the limits set by the quantities of plant, equipment, managerial and executive labour, etc., at each firm's disposal; in the long-run, the total supply of the product may be augmented or depleted by changes in both the number and size of firms. We have already seen that the ease with which these long-run adjustments can be effected is reflected in the price elasticity of the long-run supply curve of the product: if no long-run adjustment was possible, then the long-run and short-run supply curves would coincide with one another; if all long-run changes could be accomplished with perfect ease and without cost the long-run supply curve would be perfectly elastic, and following any initial rise in demand the price of the product would ultimately subside to its initial level. We shall now list the things on which the elasticity of the long-run supply curve of a product depend, and it is convenient to do so by stating the conditions that must be fulfilled if the long-run supply curve is to be perfectly elastic.

If the long-run supply curve of a product is perfectly elastic, we shall say that the industry is in a state of *perfect competition*, or that the firms in it are operating under conditions of perfect competition. Alternatively, we may say that *free* competition obtains, where 'free' means both the total absence of any restrictions of any kind on the entry of new firms or the exit of old firms, and, a consequence of this, that an increase in the output of the industry may be obtained in the long-run without any increase in the average total cost of production per unit of the product, and therefore without any rise in the price.

An economist would predict that the price of a commodity X would ultimately return to its initial level, after an initial long-run equilibrium has been upset by a rise in demand, if (*a*) there are no legal and institutional barriers in the way of new entrants to the industry, and (*b*) if the entrepreneurs that enter the industry are identical in quality with existing entrepreneurs and if they can buy productive services of the same quality and at the same price as those that are already there. For these conditions

to obtain, the following assumptions must be descriptively accurate:

(*a*) There must be no legal or institutional restrictions on the entry of new firms into the industry that produces *X*. New entry may be restricted by government or by the firms already in the industry. In the United Kingdom, for example, there is only one firm in the broadcasting, coal, electricity and rail transport industries, and the entry of new firms is prohibited by statute. The firms already in an industry may discourage new entrants by threatening to undercut their prices, by boycotting buyers who patronise them, or by collectively acquiring the source of some necessary raw material or the total supply of some essential skill and so depriving potential newcomers of its use. Existing firms may have patented the product, or some of the processes by which alone it can be produced: until the patent expires, new firms can only enter by paying licence fees to the patent-owners and these may put them in a disadvantageous position. It is only when all such obstacles are absent that new firms will be able to enter the industry and operate in it on terms that are no less favourable than those which obtain for the firms that are already there.

(*b*) If the prices of productive services are to remain unchanged as the industry expands, then each must be in perfectly elastic supply to the industry. If the productive service in question is the product of other firms, it will be in perfectly elastic long-run supply to the industry producing *X* only if the firms that produce it are themselves operating under conditions of perfect competition, so that the pre-conditions of perfect competition (that we are at present enumerating for industry *X*) must be fulfilled also in the industries producing the services that *X* buys. A productive factor (like a particular kind of labour-service) will be in perfectly elastic supply to industry *X* if the following conditions are fulfilled:

(i) Each unit of it must be perfectly mobile, both geographically and occupationally. Let us suppose that the service is carpentry, and that the demand for *X* (in whose production carpenters assist) rises. As new entrepreneurs are attracted to the production of *X*, the demand for carpenters' services will rise. If the actual (or potential) carpenters who might meet this

new demand are not living in the same places as the firms that want their services, then a higher wage-rate might have to be offered to them — a wage sufficiently higher to amortise the initial, and cover the recurrent, costs of moving. By perfect geographical mobility of a resource, we mean the absence of any money cost in transferring it from one place to another. For as many more carpenters to be forthcoming at the existing wage-rate as are required, that rate must be sufficient to cover the costs that workers with other skills would incur in changing their occupation. A high degree of occupational mobility implies the absence of any legal or institutional barriers in the way of new entrants to any trade; in addition, it requires either that the precise skills in different occupations are closely similar, or that enough workers are highly versatile and flexible in intellectual ability and manual dexterity. If there is occupational immobility, then a progressively higher wage-rate must be offered to carpenters to induce new workers to acquire that skill. What we have said of carpenters applies equally to entrepreneurs: they, too, must be perfectly willing to move, into the industry that produces X and at the rewards that can be initially earned there, from the industries and places in which they are at present engaged.

(ii) Each unit of each productive service must have full knowledge of the alternative opportunities that are open to it. Each worker, for example, must know the different jobs for which he is by natural endowment suited, and the wage-rate that might be earned in each of them. If potential carpenters were ignorant of these things, then a rise in the carpenters' wage-rate might pass unnoticed and so cause no increase in the number of carpenters. When the demand for product X rises, causing an increase in the net revenue of the firms that make it, entrepreneurs in all other industries must be aware of this fact; at least some of them must have enough knowledge of the methods and costs of production of X to realise that the higher net revenues that firms now making X are earning might be earned by them also. If entrepreneurs in other industries are ignorant of these things, then no new firms may be set up following the increase in the demand for X.

(iii) Each unit of each productive service must not only have full knowledge of present opportunities: it must also have unique

expectations about how the range of opportunities, and the reward that each promises, will vary in the future. Let us suppose that the wage-rate of carpenters rises in the short-run, as a result of a rise in the demand for their services. This need not evoke an increase in the number of carpenters, even if all the conditions that we have already listed are fulfilled, for potential carpenters may not respond because they are uncertain about the future behaviour of the carpenters' wage-rate, and because, being uncertain, they may be loath to acquire a new skill that would force them to live in the presence of uncertainty as they practised it. Thus, the range of values within which potential carpenter *A* believes the wage-rate will lie at some date in the future, or fluctuate over the future, may be wider or narrower than that which carpenter *B* has in mind, or the two ranges may overlap one another. In these circumstances, even though *A* and *B* are in all other respects identical, the wage-rate that will induce *A* to become a carpenter need not induce *B* to do so or to remain so. Different expectations about the future behaviour of the carpenters' wage-rate may, therefore, by themselves explain a less than perfectly elastic supply curve of carpenters in the long-run. Even if *A*'s expectations are identical with those of *B* — even if they both feel that the wage-rate in future will not be more than 25 per cent greater or 25 per cent less than its present level — they may differ in their attitudes towards uncertainty.* If *A* is venturous, and *B* timorous and happy only if his future earnings seem stable and secure, then *A* may move into carpentry while *B* remains where he is. What we have said of carpenters applies equally to the owners of machines and equipment and to those who are venturing their savings. Uncertainty, and the attitudes towards uncertainty, may therefore explain why the long-run equilibrium price of a productive service may rise if there is a permanent increase in the demand for it. We may think of this higher price as being a reward to those who earn it for 'bearing' uncertainty or for 'living with' it; or we may view it as being caused by the unwillingness of others to bear the uncertainty.

Uncertainty may exert its strongest influence in shaping the

* We are here assuming that the range of expected values of the wage-rate is independent of the attitude towards uncertainty. While this seems reasonable, it may not be realistic.

decision of the entrepreneur. The entrepreneur is the productive factor whose initial decision creates the firm: it is he who hires or buys productive services and organises their transformation into saleable products, and who obtains his reward in each period from the difference between the revenue and the total expenditure on all productive services. In calculating the net revenue he might earn were he to enter some industry X, he must estimate the present level and probable future behaviour of the price of the product, and of the prices and physical productivities of the productive services. Uncertainty bears more heavily on him, therefore, largely because there are more variables about whose values he may be uncertain. He may eliminate some of the uncertainty by making long-term contracts with the owners of productive services, though his inclination to do so will reflect his own attitude towards uncertainty. Even though the existing firms in industry X are currently earning high net revenues, therefore, other entrepreneurs may not enter it: they may doubt that similar net revenues would accrue to them were they to begin producing X or they may be uncertain about the behaviour of net revenue in that industry in the future. The relatively higher net revenues, that the firms now producing X are earning, may therefore be explained by the fact that the existence of uncertainty makes the supply of units of an otherwise homogeneous entrepreneurial factor less than perfectly elastic; alternatively, we may view these net revenues as being the reward that accrues to entrepreneurs already in the X-industry for 'bearing' the uncertainty.*

It is clear, then, that if each productive service is to be in perfectly elastic supply to industry X, each owner of each service must have perfect foresight about the future behaviour of the price of the service that he sells. Alternatively, we may assume that the owners of a productive service are equally uncertain about the future and what it holds for them, and that they all have the same attitude towards uncertainty. The former assumption is merely the limiting case of the latter when the 'value' of the uncertainty is zero.

(iv) Lastly, we must assume that each productive service is perfectly divisible. Let us suppose, by way of example to show the necessity for this assumption, that carpenters are an indivi-

* See the section on profit in Chapter 10.

sible factor of production. In Diagram 89, we measure the hourly wage-rate on the vertical axis; on the horizontal axis, we measure both the number of carpenters and hours of work, assuming that carpenters and those who employ them regard a working week of 40 hours as 'normal' when the wage-rate is at its long-run equilibrium level of OW. The curve S_1S_1 shows the short-run supply of hours of work when one carpenter only is employed, and similarly the curves S_2S_2 and S_3S_3. We suppose that the demand for carpenters' services is initially DD and that the wage-rate is OW: at this wage-rate, two carpenters are just willing to offer the 'normal' hours of work in each week and no actual (or potential) carpenter elsewhere feels attracted to this industry. Let us now suppose that the demand for carpenters' services rises to D_1D_1. In the ensuing short-run, the hourly wage-rate will rise to OS; this rate will appear attractive to workers in other occupations, but when a third carpenter enters, the wage-rate will fall to OL. In these circumstances, if a wage-rate of OW is sufficient to retain a third carpenter in that industry in the long-run, and if he possessed perfect foresight, he would not decide to enter the industry until the short-run rate had reached OT; for a present rate of OT is needed to ensure for him the long-

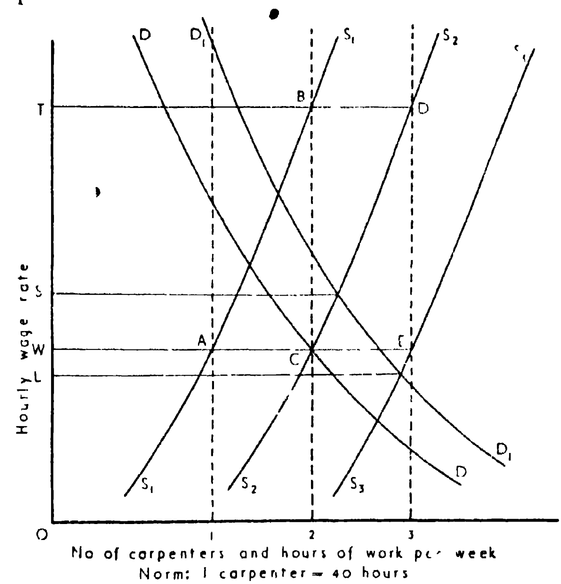


DIAGRAM 89

run rate of OW after he has actually begun work in the industry. When the demand for carpenters rises continuously, therefore, the number of hours of work that are being supplied per week will rise discontinuously along the path $WABCDE$ . . . . The existence of indivisibility — or more accurately, the fact that the quantity of a productive service cannot be increased by the same proportion (with no change in its price) as the change in the demand for it — means that the long-run supply curve of it may be less than perfectly elastic over at least a part of its range. When this occurs, we shall only observe a perfectly elastic long-run supply curve for the service if the demand for it rises discontinuously also by the same steps as the discontinuities caused by the indivisibility.

Indivisibilities may explain why there is a permanent rise in the net revenues of firms producing some product X following a rise in the demand for it, in the same way as they explain why the hourly wage-rate of carpenters may remain higher than its initial level following a rise in the demand for their services. The indivisible factor may be the entrepreneur: it may happen, for example, that the quantities of all productive services that he must employ to give him the lowest average cost per unit of output would so enhance the output of the product that its price would fall to a level that would give the new entrepreneur (and those already in the industry) a negative net revenue. Alternatively, the dominant indivisibility may lie in the machines, equipment or processes that are required in industry X: in these circumstances, the advent of a new firm using the indivisible machine or process would so enhance the industry's output that no firm earned a positive net revenue.

When there are no legal or institutional obstacles in the way of new firms entering industry X, and when all the productive services required to produce X are perfectly mobile and perfectly divisible, and their owners perfectly knowledgeable and possessed of perfect foresight, then conditions of perfect competition obtain in that industry. When these conditions prevail, then each firm producing X will be of the same size — that is, each firm will be enjoying the same advantages as each other: each firm will be employing the same quantity of each productive service as each other firm and all will be producing and

selling the same output in each period; and the total revenue being earned by each firm will just suffice to cover its total costs of production — that is, to pay for the productive services that it uses at their current market prices.

The pre-conditions of pure competition that we listed earlier define the shape of the demand curve for the product that each firm sells and for the services that each household sells: the demand curve that faces each seller of a product or productive service will be perfectly elastic at the ruling market price. The pre-conditions of perfect competition, in their turn, define the relationship between the demand and cost curves of the individual firm: the demand curve for its product and the average total cost curve will be tangential to one another, so that the firm's total revenue just covers its total costs of production. It should be noted that conditions of pure and perfect competition need not necessarily co-exist. It is conceivable, for example, that there might be pure competition in an industry, but that the entry of new firms is prohibited by government; this is unlikely, however, for government intervention is usually the consequence of organised lobbying by the firms already in the industry, and once they have enjoyed the fruits that co-operation bears they are unlikely to return to a situation in which they must act independently of one another. If perfect competition exists, however, then it is probable that pure competition exists also: for example, if all productive resources are perfectly divisible — one of the pre-conditions of perfect competition — then this by itself suggests that there will be a large number of firms already in the industry. The concepts of pure and perfect competition are, therefore, logically separate: the former defines the equilibrium of each firm in an industry; the latter defines a particular equilibrium position for the whole group of firms that constitute the industry.

If conditions of pure and perfect competition are simultaneously fulfilled in each industry in the economy, then the pattern of production will reflect exactly the pattern of households' tastes: it would be impossible for households to obtain one unit more of product *X*, for example, without their forgoing a quantity of some other product *Y* that they valued at least as highly as the additional unit of *X*. This model of universal pure and perfect competition is the 'ideal' of welfare economics. Here,

however, we are more interested in its positive than its normative uses, and we shall now use it as an 'ideal type' and deduce a system of market classification from the observed differences between the real world and the assumptions on which it (the model) rests.

A CLASSIFICATION OF MARKETS

If all the assumptions that we have stated in the previous two sections were simultaneously true of any actual market, then those operating in that market would be doing so under conditions of pure and perfect competition. The most cursory knowledge of actual markets suffices, however, to convince us that these assumptions are generally, if not always, descriptively inaccurate. While generally lacking in empirical validity, these assumptions nevertheless provide us with the criteria on which a rough and workable classification of actual markets can be based. The manner in which they help us in this respect becomes clear if we interpret each assumption as being compounded of a statement that a certain variable exercises a determining influence on the behaviour of firms, and of a statement that this variable is assumed to have a particular value. Viewed in this light, each of our assumptions becomes the 'product' of a 'multiplicand' which is simply a statement that something is relevant if we are concerned with a firm's behaviour, and of a 'multiplier' which is the precise value that is attached to this relevant variable. Thus, the assumption that to have pure competition, we must, *inter alia*, have a large number of firms in the industry, means that the behaviour of each firm (whether it will be a 'price-taker' or a 'price-maker') will depend on the number of other firms producing the product, and that to the relevant variable 'number of firms' we have attached the value of infinity. Again, the assumption that the product that is being produced by the many firms must be homogeneous, means that the 'degree of homogeneity' is a relevant variable, and that we have given it a value of infinity. Lastly, we have asserted that the behaviour of the price of a product following an increase in the demand for it will depend on the mobility and divisibility of productive resources and on the knowledge and foresight possessed by their owners and we have given each of these the value of 'perfect'.

Our judgement that the assumptions underlying pure and perfect competition are 'unrealistic' can now be expressed more precisely: each of our assumptions isolates a relevant variable, and it is the value that we have attached to each of these that is 'unrealistic'. A classificatory system into which actual markets can be fitted may therefore be developed by giving realistic and typical values to the relevant variables. The completeness of the classification that emerges will depend on whether or not all the relevant market characteristics have in fact been included in our model. The test of its completeness is partly logical and partly empirical. In our models of pure and perfect competition we obtained determinate equilibria and this suggests that nothing that was relevant was excluded. Most actual markets at the present time can be brought within the classification by varying the values that we attach to the characteristics that we have isolated. Lastly, the success with which we can explain economic events and predict their main consequences with the help of hypotheses derived from these assumptions suggests that all the major influences that affect the behaviour of relative prices have been included.

Before proceeding to list typical markets, it is worth noticing the analogy between our intentions as described in the previous paragraphs and the methods of the physical sciences. The law of gravity, for example, states that if a body of any size, shape or mass is dropped near the earth and in a vacuum, it will fall with a constant acceleration of about 32 feet per second per second. This is a limiting case of a more general statement, namely, that the rate of fall of a body depends on its size, shape and mass, the height from which it is dropped and on the atmospheric pressure: the law gives the value of zero to the last of these, making size, shape and mass of no importance, and it gives a low value to the height from which the body is dropped. By giving other and more 'realistic' values to these variables, we could derive alternative hypotheses about the rate of acceleration, and we would find that there is close correspondence between the values accorded to the variables and the extent of the divergence of the rate of acceleration from 32 feet per second per second. By proceeding in that way, we can make more precise predictions about the behaviour of particular objects dropped in particular circumstances. The law of gravity is analogous to the hypotheses

that have been put forward in the previous eight chapters. These hypotheses enabled us to predict the general direction of the changes in relative prices that would follow a particular economic event. As we give more realistic values to the variables relevant to our analysis, we may be able not merely to predict the general direction of the change in the relative price of a commodity following a change in the demand for it, but also to give some notion of its extent by indicating whether it would be less or greater than that which would have occurred had conditions of pure and/or perfect competition prevailed.

The variables whose value helps to determine the actual course of particular prices are (i) the number of sellers; (ii) the number of buyers; (iii) the amount of knowledge that each possesses; (iv) their objectives; (v) the degree of homogeneity of the product they buy and sell; (vi) the unimportance of legal and institutional barriers to the entry of new buyers and sellers; (vii) the mobility and (viii) the divisibility of productive resources; (ix) the degree of knowledge and (x) of foresight possessed by their owners. Each of these may assume any value from zero to infinity, and since we believe that this list includes all the relevant variables, there must exist a set of precise values for each of them that will accurately describe the morphology of any particular market in the economy. The infinite number of possible market morphologies that may emerge in that way has customarily been fitted into a primary classification that arises in the following manner: a value of zero is given to variables (vi) to (x) inclusive, and variables (iii) to (v) inclusive are given the same values as under pure competition, and alternative values are given to the numbers of buyers and sellers of the product.

Economists have long been aware that a high degree of correlation exists between the number of sellers (or of buyers) of a commodity and the market behaviour of each of them, so that the values that were accorded to these variables were the critical values necessary to isolate the different kinds of market behaviour. On this basis, the following classification emerged:

1. *Monopoly:* This is the name given to the market form in which there is one seller, an infinitely large number of buyers, and in which the values assumed by variables (iii) to (v) are the same as when conditions of pure competition exist, and the

value of each of the variables (vi) to (x) inclusive is zero. Where this kind of market exists, we would expect the single seller to be an 'independent price-maker' — that is, to possess some power to determine the price of his product — and we generally find that these expectations are confirmed. The complementary market form is *monopsony*, where there is one buyer, an infinitely large number of sellers, and where all the other variables have the same values as for monopoly.

2. *Oligopoly:* This is a market in which there are a few sellers, and in which the values assumed by all the other variables are the same as when there is monopoly. When oligopoly exists, we would expect each of the small number of sellers to have some power to choose the price at which he will sell his product, but this power is limited by the existence of a few other firms selling the same product. For brevity's sake, we shall say that each oligopolist is an 'interdependent price-maker' — that is, his power to set a price for his product is circumscribed by the decisions of his rivals. The complementary type of market is *oligopsony*, where we have a few buyers and an infinitely large number of sellers, and where all the other variables have the same values as for oligopoly.

3. *Bilateral Monopoly:* Here, there is one buyer and one seller, and each of the other relevant variables has the same value as for monopoly or oligopoly. When bilateral monopoly exists, the single buyer and the single seller of the commodity or productive service each possesses some power to fix the price at which the transactions between them shall be finally effected.

4. *Monopolistic Competition:* This form of market was first explicitly isolated by Professor E. H. Chamberlin in his *The Theory of Monopolistic Competition* (Harvard University Press, 1933), so that it is the youngest of this family of primary types of market. A state of monopolistic competition exists when all the variables (save (v) and (vi)) are given the same values as under conditions of pure and perfect competition. It is assumed that each unit of the product which is being produced by many firms and bought by many households is a close, but not a perfect, substitute for each other unit, and that there is some legal obstacle (such as the laws relating to patents and trade-marks) that prevents any firm from producing and selling a product that is in *all* respects identical with that being currently offered by any other firm.

If we wish to make this list of market prototypes exhaustive, we must add the market forms with which we are already familiar:

5. *Pure Competition:* In its simplest form, this type of market requires the values for variables (i) to (v) that we have listed on pages 261–5 above, and zero values for the variables (vi) to (x).
6. *Perfect Competition:* This type of market logically requires that pure competition exists also, so that for it to occur we must give the pure competition values to (i) to (v) and the values of infinity to variables (vi) to (x).

In the chapters which follow, we shall concentrate mainly on the kinds of competition that are to be found in the markets in which products are bought and sold. In the next chapter, we shall study the prototypes of monopoly/monopsony and monopolistic competition, and we shall examine the commoner variants of each of them. In the following chapter — Chapter 11 — we shall study oligopoly/oligopsony and bilateral monopoly.

CHAPTER 10

Monopoly and Monopolistic Competition

The manner in which capitalistic economic systems have evolved has resulted in products and productive services being produced, bought and sold in a wide variety of different market structures. In this chapter and in the next, we shall describe the way in which these markets are customarily classified. In any system of classification, objects that are put into the same class are described by listing some of the attributes that they have in common, and the attributes that are actually listed in the definition of the class are those which are believed to be the most important. In economics, markets are classified according to the number of buyers and sellers who operate in them and the degree of homogeneity of the product that they buy and sell. In all the markets that we collectively call 'monopoly', for example, there is one seller and many buyers of a homogeneous commodity; in the class that we call 'oligopoly', there are a few sellers and many buyers of a commodity that is either perfectly homogeneous, or sufficiently homogeneous for it to be meaningful to call each unit of it by the same name. Each class has many other attributes in addition to those we list in our description of the markets that are included in it: for example, the degree of freedom that a seller possesses in deciding at what price to sell his product varies widely from one individual market to another. This, however, is regarded as a subordinate attribute, in that variations in a firm's freedom to fix a price for its product seem to be closely correlated with the number of firms that are currently producing the same (or closely similar) product and with the number of households that are currently buying it. We shall therefore find that our classification which is based on the morphology of markets accords closely with a classification based on the market behaviour of firms. In this and in the next chapter, we shall describe the market behaviour that is typical — and indeed, probably the necessary consequence — of the classes of

market that we have called monopoly/monopsony, oligopoly/oligopsony, and bilateral monopoly. This can be done most easily by taking an 'ideal type' for each class of market and studying its properties. In no case, however, is it likely that all (or indeed any) of the actual markets, that are put into any of our classes, will conform to the 'ideal type' that we choose, for we would expect to find as great differences between markets that we group together as monopolies or oligopolies as are found between individual beings that are grouped together in the biological class *homo sapiens*. We shall therefore indicate the kind of variations that may be found between the individual members of each of our main classes. Lastly, when all that has been done, we shall attempt to assess the usefulness of the system of market classification that we have adopted, and of the theories of market behaviour that have been based on it.

MONOPOLY

We shall take the simplest case of monopoly for our 'ideal type'. Let us suppose that there is one seller of commodity *X*, that pure competition exists in the markets in which he buys his productive services, so that the price of each of them is a datum for him, and that there is a very large number of knowledgeable buyers who buy his product in each period of time. We shall further suppose that the monopolist believes that there is no possibility, either now or in the forseeable future, of any new firm(s) being set up to produce *X*, and that, in pursuing his objective of earning the maximum net revenue per period, he believes that his actions do not affect in any way the prices of any other products or the behaviour of the firms that make and sell them. By these assumptions, we, *inter alia*, exclude monopsony in the markets in which the monopolist buys his productive services and we eliminate all elements of oligopoly in the market in which he sells his product. Starting with this simple model of monopoly, we shall attempt to do three things: first, to describe the typical market behaviour of a monopolist; second, to indicate the wide variety that may exist amongst the individual markets that are classified together as monopolies and this we shall do by modifying some of the assumptions on which the simple model of monopoly rests and examining the consequent modifications in the monopolist's plans; and third,

to catalogue the methods by which monopoly might be created and the measures by which it might be perpetuated.

THE TYPICAL BEHAVIOUR OF A MONOPOLIST

Our monopolist is the sole seller of X. There are many knowledgeable buyers of X, and the relationship between their planned purchases of X and its price is summarised in the total demand for X; and we know (from Chapter 1) that the planned purchases at each price in each period of time depend on the tastes and preferences of the buyers, their planned consumption expenditures and the price of each other commodity that competes with X for their patronage. The total demand for X, therefore, delimits the sales environment of the monopolist—it describes the whole range of sales plans that is open to him. He can, if he so desires, plan to sell a relatively small quantity of X at a relatively high price, or a relatively large quantity at a relatively low price. The price that he decides to fix for his product. and therefore the quantity of it that he can hope to sell, will depend on his objective. In this simple model, we are assuming that he seeks to earn the maximum net revenue per period, so that the monopolist will choose the sales plan, and the purchase plan for productive services by which it can most cheaply be implemented, which promise the maximum excess of total revenue over total costs. This choice is illustrated in Diagram 90, where we assume that the monopolist is fully aware of the demand for his product and of the relationship between total costs of production and output.

The demand curve for X is shown by DD in Diagram 90(a). The relationship between sales and total revenue, which is implicit in this demand curve, is shown by the curve OR in Diagram 90(b). Each point on the OR-curve may be derived simply from the demand curve: thus, an output of OM can be sold at a price of OP per unit, giving a total revenue of $OMNP$, which is represented by the distance MS in Diagram 90(b). The behaviour of the OR-curve reflects the price elasticity of the demand for X: to effect an increase in the sales of X per period the price must be lowered, and recalling the approximate measure of the price elasticity described in Chapter 1 (*supra*, pages 37–9) we see that the revenue curve will rise while the demand is relatively elastic, reach a maximum when demand has unit elas-

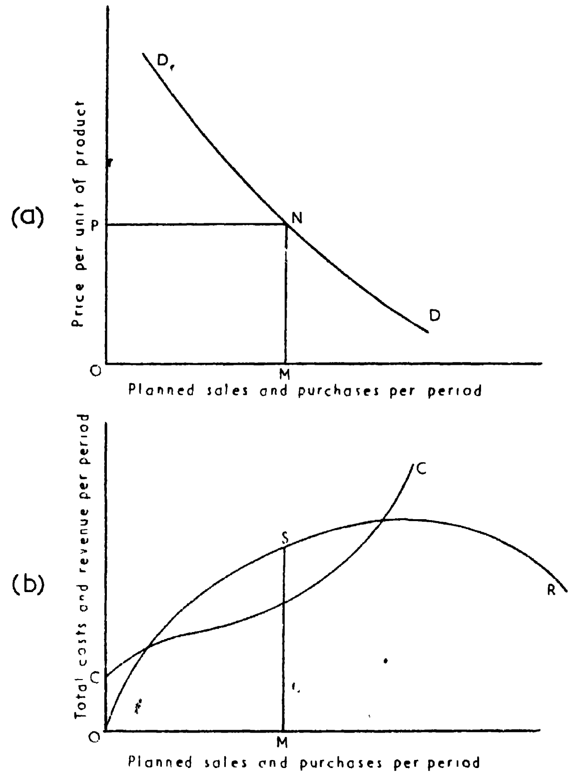


DIAGRAM 90

ticity and fall when demand is relatively inelastic. This relationship between revenue and the price elasticity of demand may be expressed precisely: let dR be the small change in revenue following a small change in sales of dX that is effected by a small reduction in the unit price of dP. Then, if initially the price was P per unit and sales X per period:

dR = Total revenue after the price reduction *less* total revenue before the price reduction

$$= (P - dP)(X + dX) - P \, . \, X$$

$$= P \, . \, X - X \, . \, dP + P \, . \, dX - dP \, . \, dX - P \, . \, X.$$

That is:

$$dR = P \, . \, dX - X \, . \, dP \text{ (approximately)}$$

$$= P \, . \, dX(1 - X \, . \, dP/P \, . \, dX).$$

We know, however, that $X \,.\, dP/P \,.\, dX$ is equal to the reciprocal of the price elasticity of demand, ϵ. Hence

$$dR = P \,.\, dX(1 - 1/\epsilon)$$

or

$$dR/dX = P(1 - 1/e).$$

The expression, dR/dX, is called the 'marginal revenue', and we shall deal with it more fully later.

The line CC in Diagram 90(b) shows the relationship between the monopolist's total costs and his rate of output, and it may be obtained by the methods described in Chapter 2. The OR-curve shows us the total revenue that he could earn by selling each output, and the CC-curve shows us the minimum sums of money that he must spend on hiring or buying the productive services that are needed to produce each output. The net revenue that he would earn by producing and selling any output is shown by the vertical distance between the revenue and cost curves at that output. Given the objective that we have imputed to him, he will plan to produce and sell OM in each period at a unit price of MS/OM, for then, and only then, will his net revenue per period be at a maximum.

The monopolist's choice of a sales plan may be described in different words. If net revenue is at a maximum at the output OM, then it must be less than this at outputs which are greater or less than OM by one unit: that is, costs must rise faster than revenue as output is expanded beyond OM, and revenue must fall faster than costs as output is reduced below OM. At the output OM, therefore, the rate of change of revenue must be the same as the rate of change of total costs. We know that the rate of change of total costs as output is varied is called marginal cost (*supra*, Chapter 2, page 74) and that the marginal cost at any output is measured by the slope of the total cost curve. The rate of change of revenue as sales are varied is called *marginal revenue*, and it is measured by the slope of the total revenue curve at each level of sales. From Diagram 90(b), we can see that the slope of the total revenue curve continuously declines as sales are increased: it becomes zero when revenue is at a maximum and assumes progressively larger negative values thereafter. We can see also from Diagram 90(b) that at each output marginal revenue must always be less than the unit price

at which that output can be sold: thus, at the output OM, marginal revenue is equal to the slope of the straight line which is a tangent to the total revenue curve at S, and the price or average revenue is equal to the slope of the straight line OS, and it is clear that the latter rises more steeply (that is, is greater) than the former.

The behaviour of the marginal revenue, and its general relation to the demand or average revenue curve, is shown by the line labelled MR in Diagram 91. If on this diagram we draw also the monopolist's average total and marginal costs of production (AC and MC respectively), we can illustrate his choice of a sales plan, for the output that promises him the maximum net revenue in each period will be that at which marginal cost and marginal revenue are equal to one another. The monopolist will therefore plan to produce an output of OM per period — the output at which the marginal cost and marginal revenue curves intersect one another — and sell it at a price of OP per unit to earn a maximum net revenue of $RTWP$ per period. Since the demand curve in Diagram 91 is identical with that drawn in Diagram 90(a), the output of OM in the former will be identical with the output of OM in Diagram 90(b). We saw earlier that at any price marginal revenue equals $P(1 - 1/e)$; if

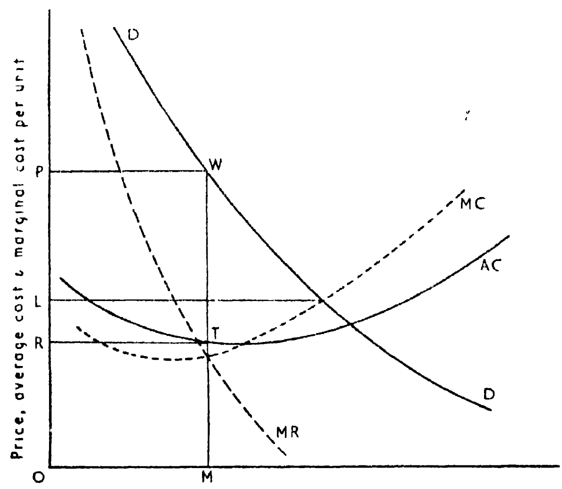


DIAGRAM 91

the price is also that which promises the maximum net revenue per period, then marginal cost and marginal revenue must be the same; so that $MR = MC = P(1 - 1/e)$, or $P = MC/(1 - 1/e)$. In developing our analysis of monopoly, we shall generally use the curves illustrated in Diagram 91. It must be emphasised, however, that when we say that the monopolist chooses the output at which marginal cost and marginal revenue are equal to one another, we add nothing to the statement that he chooses the output which promises him the greatest net revenue per period.

SOME ASSUMPTIONS MODIFIED

In this section, we shall modify some of the assumptions that have underlain our analysis of monopoly so far. We shall examine what happens when (*a*) the monopolist pursues some objective other than the maximisation of his net revenue; (*b*) there is not one market, but rather several markets, in which he may sell his product; (*c*) all potential buyers are not aware of the monopolist's product or of its relative ability to satisfy their desires, and (*d*) the monopolist believes that the demand for his product in future depends on his present actions. First, the alternative objectives that a monopolist might pursue. In our example, the fact that the monopolist fixed the price of his product at *OP* per unit was explained (given all the other assumptions) by his assumed desire to maximise his net revenue, in each period of time. Now this objective is the same as that of a firm operating under conditions of perfect competition. For a firm in those conditions, however, that objective is obligatory: it is not so much a separate element in the market morphology as a necessary consequence of all the other elements, for if a firm is to survive in a perfectly competitive market it can only hope to succeed in doing so by seeking the maximum net revenue from its current operations. A firm that is operating under conditions of otherwise pure competition might regard this objective as permissive in the short-run, but if competition is also perfect it must pursue it to avoid bankruptcy in the long-run. Our monopolist is not compelled, however, to choose this objective, for there is no other firm, either now or in the future, to so compel him. The assumptions by which we have defined the environment in which he operates will not help us, then, to say what particular aim he will have, for that will depend largely on the individual idiosyn-

cracies of the person who plays the role of monopolist in our model. While in these circumstances, our model will not help us to predict the precise price and output that will be fixed in any particular case, it nevertheless helps us to delimit the range of highly probable prices and outputs: the monopolist will not fix a price that lies below that at which the average cost and demand curves intersect one another in Diagram 91, for at prices below that his total revenue would not cover his total costs, and he is unlikely to fix his price above OP, the price at which the net revenue per period would be at a maximum.

Second, in our analysis of monopoly so far we have assumed, *inter alia*, that the households who bought his product were as knowledgeable as they would have been had conditions of pure and perfect competition obtained. A necessary consequence of this assumption was that each unit of the monopolist's output would be sold at the same price: if he fixed a higher price for group A of buyers than for group B, the former would buy from the latter and the latter alone would buy from him; in these circumstances, since all units of the product must be sold at the same price, the monopolist will earn the maximum net revenue by fixing the price of each unit at OP in Diagram 91. We shall now suppose that the potential buyers of the product do not constitute a single and knowledgeable group, but that they are divided into several groups, and that each buyer is aware of the prices at which any other buyer in the same group is buying, but quite unaware of (or unable or unwilling to profit by) the price at which any buyer in any other group is buying. Each such group of buyers will then constitute a separate and independent market for the monopolist's product. The markets may be separated by ignorance, or by laws that prohibit the movement of the commodity from one market to another, or by accepted social conventions and customs that frown upon transactions between members of different groups. Our analysis may easily be extended to encompass several markets. The monopolist remains the sole producer and seller of X to all markets, and there will be a separate demand for his product in each market, depending on the tastes and planned expenditures of the buyers who buy within it and on the prices of all the other goods that they might buy. The monopolist's problem is to fix a price for his commodity in each market, and its formal solution is illus-

trated in Diagram 92, where we assume that there are two markets (A and B), that the monopolist knows the demand for his product in each of them and the relationship between his total costs of production and his rate of output, that his objective is to earn the maximum net revenue per period and that the cost of transporting his product to either market is zero.

The demand curve for the product in A is shown by D_aD_a in Diagram 92(a), and that in B by D_bD_b in Diagram 92(b). The monopolist's marginal costs of production are shown by the MC-curve in Diagram 92(c). When the monopolist is earning the maximum net revenue per period from the production of his product and its sale in markets A and B, he will be producing the output at which marginal cost and marginal revenue are equal to one another, for that is merely another way of saying that net revenue is maximised, and he will be distributing his output between the two markets in such a way that the last unit sold in each market adds the same sum to his total revenue, for if that is not so then a higher total revenue can be obtained from the sale of the same output by transferring sales from A to B, or vice versa. We can discover the output and price for each market at which both these conditions will be fulfilled with the help of Diagram 92, though any actual monopolist will seek the solution experimentally rather than geometrically. The marginal revenue curves of markets A and B are added together to give us the curve MR_{a+b} in Diagram 92(c). This curve shows the maximum total revenue that can be obtained by selling any

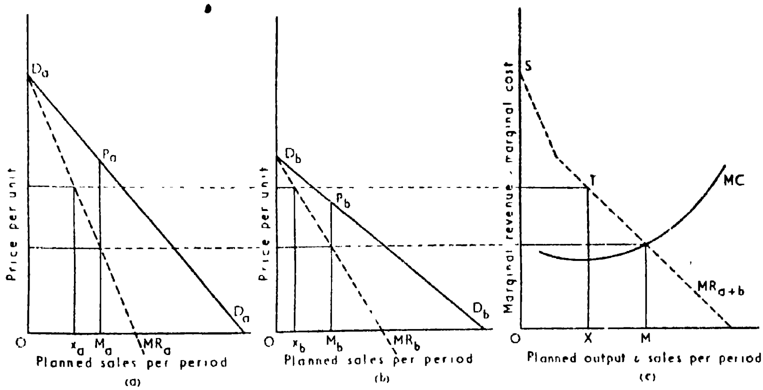


DIAGRAM 92

output: thus, $OXTS$ is the greatest sum of money that can be earned by selling an output of OX, and this is earned when a quantity OX_a is sold in A and OX_b in B.* The MR_{a+b}-curve thus illustrates the second condition mentioned above. The first condition is fulfilled when the monopolist is producing an output of OM per period, for at that output marginal cost and the addition to revenue from selling the OM-th unit are the same. The monopolist will therefore plan to produce OM per period and to sell OM_a in market A at a price of M_aP_a per unit and OM_b in B at a price of M_bP_b per unit.

We know (see *supra*, page 283) that at any level of sales in any market marginal revenue $= P(1 - 1/e)$, or $P = MR(e/e - 1)$. If the demands in markets A and B are such that at each price the price elasticity of demand is less in A than in B, then the marginal revenue yielded by selling an additional unit in A will be less than the marginal revenue in B. Conversely, if the marginal revenues are the same in A and B, then the price must be greater in A than in B. If at each price the price elasticity of demand in A is the same as that in B — that is, if the demand curves are iso-elastic — then when the marginal revenues in the two markets are the same, the monopolist will be charging the same price in both markets. We conclude, then, that if the demand curves in the separate markets A and B are iso-elastic, the monopolist will earn the maximum net revenue by charging the same price in each market — that is, there will be no price discrimination between them, for if different prices were charged the net revenue would not be maximised; if the demand curves are not iso-elastic, then when net revenue is greatest the price will be higher in the market where the price elasticity of demand (at each price) is lower.

We shall not explore the mechanics of price discrimination further.† Where there exist several independent markets for a monopolist's product, he may earn a larger net revenue by charging a different price in each of them than by lumping them together and treating them as a single market. As a corollary, if the practice of price discrimination promises a higher

* Since the MR_{a+b}-curve was derived by adding together the sales in each market at each level of marginal revenue, OX_a plus OX_b must be equal to OX.

† The classic treatment, not only of the mechanics, but also of all aspects of price discrimination is to be found in Joan Robinson: *Economics of Imperfect Competition*, Book V, London, Macmillan,

rate of net revenue, a monopolist who faces a single market for his product will have an incentive to divide it into separate markets, and we would expect him to attempt to do so provided the expected costs of dividing the market did not outweigh the expected gains from exploiting the divided market: that is, a monopolist may behave not only as a 'price-maker' but also as a 'market-divider'. Generally, the total revenue that he can earn from the sale of any given output can always be increased if the market is divided into sectors, in each of which the price elasticity of demand is different at each price from that in each other, and if he can effectively prevent the movement of his commodity from one sector to another. The division of the market may be effected in various ways. The monopolist, for example, may persuade the political authority in the part of the market where the demand is relatively inelastic to impose duties on the importation of his product from the part(s) of the market in which he plans to sell it at a lower price because the demand is there relatively elastic. Again, it may be that the elasticity of demand varies according to the use to which the product is put, and that a unit of it that is bought for one use cannot be resold to a buyer that wants it for another: thus, to take an approximate example, a railway company might vary its charges per ton-mile according to the value of the commodity that it is asked to transport. Lastly, a monopolist might effectively divide the market if he successfully convinces the buyers whose demand is relatively inelastic (and who pay a relatively high price) that the product they are buying is not the same as that which is being bought by buyers whose demands are relatively elastic (and who are therefore paying a lower price). The American Featherbed and Pillow Company may once have provided an example of an attempt to divide a market by this method: the Federal Trade Commission found that this company marketed its product under five brand names, namely, 'Princess', 'Progress', 'Washington', 'Puritan' and 'Ideal', and its advertising suggested that these products were of different qualities; the Commission found, however, that all five brands were of the same quality, and that it was the label alone that differentiated one brand from another.*

Third, we shall study the behaviour of the monopolist when he knows or suspects that all potential buyers are not aware of his product or of its relative ability to satisfy their desires. In these circumstances, the monopolist can increase the demand for his product by calling its existence and properties to the attention of all potential buyers by advertisement. He may, indeed, go further and attempt not only to increase the knowledge of buyers so that their existing tastes and preferences may be more fully satisfied, but also to intensify their preferences for his product. For a monopolist in this position, the demand curve for his product is not a datum (as it was for the monopolist in our simple example) but a variable whose value is at least partly dependent on his own actions. We shall not attempt to represent diagrammatically the choice of a sales plan by a monopolist who advertises, but shall rest content with delineating the range of choice that faces him. Each sum of money that he contemplates spending on advertisement may be spent in an infinite number of different ways, and the effect of its expenditure on the position and shape of the demand curve will depend on the way in which it is spent. Thus, a sum of £1,000 per period may be used to buy space in weekly journals or in daily newspapers, or it might be spent on handbills or posters, neon signs or sky-writing, or it might be used to pay the wages of salesmen who hawk the product from door to door. If spent on newspaper advertising, there may be a whole-page advertisement in one issue of a national daily, or a smaller advertisement in a number of successive issues. There will be a different change in demand for each way in which this sum is spent. There may be an increase in the planned purchases at each price as in Diagrams 93(*a*), (*b*) and (*c*), or an increase in planned purchases at some prices as in Diagrams 93(*d*) and (*e*); in (*a*) and (*b*) the new demand curve is less elastic than the old one at each price, and in (*d*) it is more elastic at lower prices and in (*e*) less elastic at higher prices. For each new demand curve that he might have by spending £1,000 per period on advertising, the monopolist can calculate the price and output that promises him the maximum excess of total revenue over total *production* costs,* and from this he must deduct the £1,000 he spends on advertisement

* We shall assume, for simplicity's sake, that there is no change in the physical characteristics of the product so that the costs of production remain unaltered.

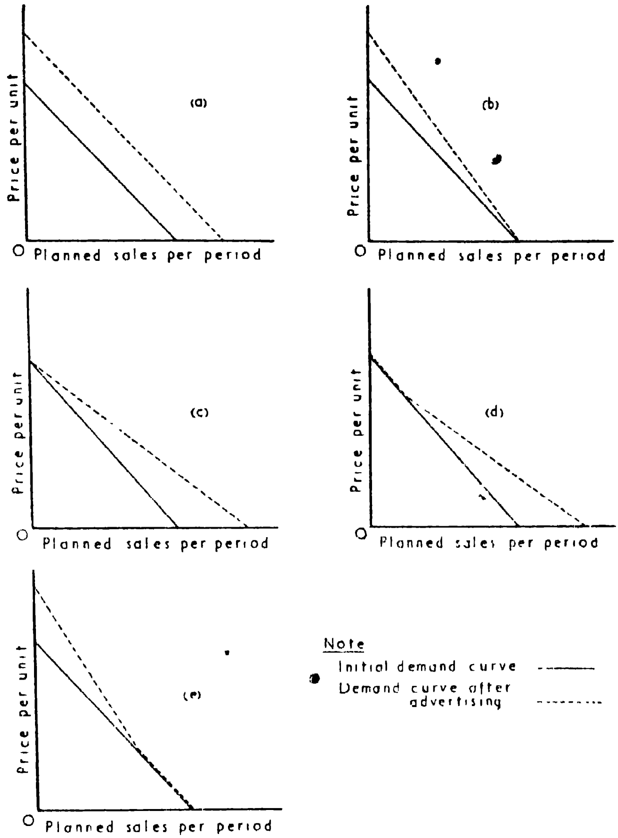


DIAGRAM 93

to get his net revenue. A similar calculation can be made for each other level of advertising expenditure. For each sum of money that he spends on advertising, there will be a particular way of spending it that promises the greatest net revenue. From all these maximum net revenues he will choose the *maximum maximorum*, and in doing so, he will be simultaneously fixing the price of his product, the output that he will produce in each period, the level of advertising expenditure and the manner in which to spend it.

In the previous paragraph we have assumed that the product remains the same for the monopolist, and that he attempts to earn a higher net revenue through expanding his sales of the

given product by advertisement. It may be possible, however, to increase his net revenue per period by altering the design, colour, packaging, or any other attribute of the product. The mechanics of this decision may be indicated briefly. The costs of production and the demand will vary from one variant of the product to another. For each possible variant of the product, he can calculate the maximum net revenue that he would earn per period by producing and selling it, and in choosing the *maximum maximorum* net revenue he will be simultaneously choosing the product-variant, the price at which to sell it and the quantity to produce of it in each period.*

We have already seen that a monopolist is always a 'price-maker', and that it may be sometimes possible for him to enhance his net revenue by dividing those who buy his product into non-competing groups. We see now that he may influence the size of the market for his product, and that it may be profitable for him to do so, because the fruits of his selling efforts will accrue wholly or mainly to him since he is the sole producer of the advertised product.

Fourth, we shall examine how the monopolist's aims and behaviour might be modified if he believes both that new firms might be formed to produce the same or a closely similar product and that whether or not they will actually be set up depends on his present actions. Assuming that there are no legal or institutional barriers that effectively prevent new entry, entrepreneurs will be attracted into this field by the prospect of earning a higher net revenue per period.† They may base their belief that they could enhance their net revenues per period by competing with the monopolist on (*a*) the present price of the commodity; (*b*) the size of the net revenue that the monopolist is now earning; (*c*) a feeling that the monopolist is not effectively catering for the market for his product, either because many potential buyers remain ignorant of its existence or because some or all buyers would prefer a variant of the product that the monopolist appears unwilling to offer them; and (*d*) a

* For a formal treatment of the monopolist's decision when price/output, product-quality and advertising expenditure are all variable, see Hans Brems: 'The Interdependence of Quality Variations, Selling Efforts and Price', *Quarterly Journal of Economics*, May 1948.

† Strictly, by a rate of net revenue that exceeds the rate they are at present earning by enough to cover the expected costs of moving.

belief that they could produce the monopolist's product more cheaply, either by using newer techniques of production or by more efficient organisation and management within the technique that the monopolist is now using. If the monopolist is aware that these are the criteria on which potential competitors will base their decisions, and if he wishes to retain the whole market for himself over the long-run, then he will attempt to fix values for the price of his product, his net revenue per period, his advertising expenditure and technique of production that effectively discourage new entry. We may call his objective in these circumstances the maximisation of the 'present value' of the stream of net revenues per period over the long-run — that is, of the sum of the expected net revenue in each future period discounted to a present value at what is for him the relevant rate of interest. It is not possible to indicate with any degree of precision the value that the monopolist must give to his price, net revenue, advertising expenditure and costs of production if potential competitors are to be permanently discouraged. All that we can say is that a monopolist who seeks to remain a monopolist will pay more attention to the magnitude and method of his advertising and will experience a stronger urge to improve the techniques by which his product might be produced and so lower its costs of production. Having done all this, if he feels that new entry still threatens, he will reduce his price below the level that promises the maximum net revenue in the current conditions of demand for the product, and so lower his present net revenue: we can say little more than that the price which he fixes will be less than *OP* but greater than *OR* in Diagram 94. Its precise level will depend on his estimates of what the production costs of new firms would be were they to appear and of the rate of net revenue that would just induce them to enter.

A monopolist who is not harassed by the threat of competition may fear that his customers will judge both his price and his net revenue 'too high' if he seeks the maximum net revenue per period, and that they might succeed in having his operations curbed or controlled by the political authority. The monopolist will respond to the prospect of potential control of his price and net revenue in the same way as he does to that of potential competition. If he believes that either competition or control is inevitable, and that it will occur despite any effort he might make to

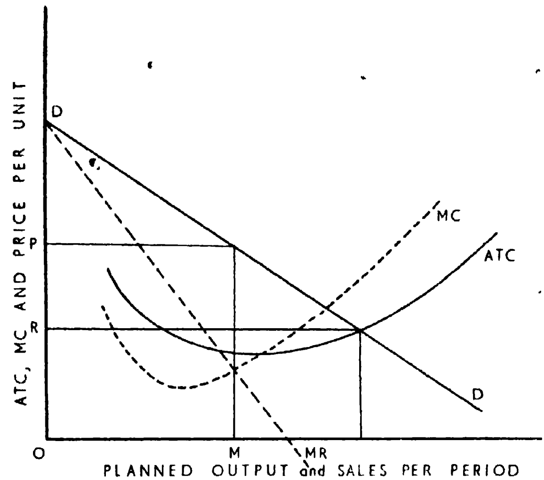


DIAGRAM 94

prevent it, then his present aims and behaviour will not be altered: he will continue to extract the maximum net revenue from the market for his product while he can, for by doing so he is not jeopardising his net revenues in the future.

GENESIS OF MONOPOLY AND DEVICES FOR PERPETUATING MONOPOLY

Monopoly may be a natural consequence of the fact that (for some commodities and services) the unit costs of production are lower for large than for small outputs, or it may be the result of conscious efforts directed towards establishing it. The output in each period from the plant that gives the lowest average total costs per unit of the product may be large enough to meet the planned purchases of buyers at all prices at which it is likely to be sold. If more than one such plant existed, some or all of them would earn a negative net revenue and thus be driven into bankruptcy. Monopolies that arise for this reason are called 'natural' monopolies, and the industries supplying water, gas, electricity and rail transport are typical examples. In most countries, these natural monopolies are nationalised, municipalised or subjected to rather strict control by the government.

While some firms may have monopoly thrust upon them by the current pattern of relative prices and state of the technical

arts, it is probable that most monopolies are the result of deliberate and purposive effort. The independent firms producing a commodity may merge together to form a single firm that thereafter is the sole producer and seller; or one firm may either acquire control of all the others or drive them out of business; or the firms while preserving their separate identities as producers may agree to act in concert as sellers. In the recent past, such efforts to establish monopoly have frequently enjoyed the blessing, if not the active support, of governments. While the methods by which monopoly may be established are legion, their objective is generally the acquisition of power. When there are many independent sellers of a commodity, the power of any one of them to fix a selling price for his output is effectively circumscribed by the existence of all the others; when there is a single seller (or group of sellers acting in concert) the power to fix the price is limited only by the conditions of demand for his product. The power that monopoly confers may be sought, then, because of the higher rate of net revenue that can be earned by its possessors. It may be sought also to enhance the bargaining strength of those possessing it *vis-à-vis* the government or another monopolist (for example, a trade union) and thus to maintain or to increase their net revenues.

The gains that currently accrue to the monopolist wholly depend on his position as the sole seller of the product; if they are to be his permanently then his position as sole seller must be assured by the effective prevention of new entry. The market for the product of a natural monopolist is protected in the long-run by 'indivisibilities' of productive processes and factors; even here, however, the protection is not absolute for the invention of new substitutes for his product or the development of new techniques by which relatively small outputs may be produced at a unit cost as low (or lower) than that which he is now incurring may expose him to competition from new firms. A monopoly that is formed by merger, combination or agreement may enjoy no such 'natural' protection, and if it is to remain as the sole seller of the product the entry of new firms to compete with it must be prevented either by law or by its own actions. A government may protect the national market of a monopolist by imposing tariffs on the same or similar products imported from other countries. There may be legal restrictions on new entry:

thus, the licensing system set up by the Road Traffic Act of 1930 and the Road and Rail Traffic Act of 1933 restricted the entry of new firms into road transport in Great Britain. The monopolist may have patented his product or some of the processes by which it is produced so that any firm desiring to compete with him must pay him royalties or licence fees and thus suffer higher costs of production. A monopolist may deprive new entrants of markets for their output or of sources of supply of basic raw materials. He may do the former by making long-term contracts with his customers, by offering them substantial rebates that depend either on their buying solely from him or on the quantities of his product that they buy, or he may attempt to bind his customers wholly to his product by substantial and sustained expenditure on advertising. He may do the latter by making long-term contracts with the firms that supply him or by buying these firms and so assuring their output permanently to himself. To the extent that a monopolist indulges in these practices, the costs that a new firm must incur if it is to compete effectively with him are increased, and they may be made so large that new entry is prevented in practice. Lastly, new firms may be deterred from setting up to compete with the monopolist by the fear that he will drive them into bankruptcy before they are established: he might do this by depriving them of customers by deliberate price-cutting or by bribing and coercing their employees and suppliers.

In this volume we are concerned rather with description and analysis than with passing judgement on the desirability or otherwise of the phenomena that are described and analysed. If we should wish to pass judgements on monopoly, however, we may base them either on the extent to which its results deviate from some ideal (such as pure and perfect competition) or on the methods by which it is established and maintained. From a comparison of monopoly with pure and perfect competition, we can make quantitative judgements: for example, when there is only one seller of a commodity, its price will be higher and the output per period lower than would have been the case had it been produced and sold under conditions of pure and perfect competition; and if we equate 'more' with 'better' and 'less' with 'worse' then monopoly may be deemed undesirable, and therefore the practices by which it is maintained and established may

be deemed undesirable also. Alternatively, we may pass our primary judgement on the methods and practices by which the monopolist assures his position as the sole seller of his product: if we believe that each household *should* have substantial freedom to choose from whom it will buy and that each firm *should* have a wide freedom to decide what product(s) to produce and sell, then monopoly, irrespective of its results, must be deemed undesirable, for the devices by which it is maintained limit the freedom of choice of households and firms.

MONOPOLISTIC COMPETITION*

As our 'ideal type', we shall take the simplest case of monopolistic competition. We shall suppose (*a*) that there is a very large number of independent sellers of some class of commodity (like tea, motor-cars or toothpaste); (*b*) that the product of any seller is an equally close substitute for that of any other seller and that the products of all sellers are sufficiently alike to be called by the same class-name, such as motor-cars or toothpaste; (*c*) that all productive services (including the services of entrepreneurs) are in perfectly elastic supply to the production of this class of commodity, and (*d*) that there is a large number of knowledgeable buyers of the class of product that the firms are selling. We shall further suppose (*e*) that in the long-run, competition is perfect except in that no firm (new or old) may decide to produce and sell a product that is a perfect substitute for a product that is being currently offered by any other seller. Given these assumptions, there will be a separate demand curve for the product of each seller, showing the quantity of his product that buyers would plan to buy at each price in each period, given their tastes and preferences, planned consumption expenditures, and the price (*inter alia*) that is being charged by each other seller for his product. The demand curve for each firm's product will be highly elastic at each price, because there exist many close substitutes for it. Furthermore, the demand curve for the product of any seller will be independent of his own behaviour: since he is only one of a very large number of sellers

* The term 'monopolistic competition' has been given many meanings. It is probable that Professor E. H. Chamberlin used it to mean the market structures in which a differentiated product is produced either by a small or by a large number of independent firms. Since the term 'oligopoly' (suitably qualified) is available for the former, we confine 'monopolistic competition' to the latter.

and since his product is an equally close substitute for that of any other seller, if he lowers his price his gain in sales will be distributed more or less equally over all the other firms so that the extent to which any other firm suffers will be negligible, and, being negligible, will evoke no change in the price at which it is currently selling its product.

In these circumstances, each seller will have some choice in fixing the selling price of his product. This choice is illustrated in Diagram 95, where we assume that the firm is fully aware of the demand for its product and of the relationship between the total costs of production and its rate of output. The demand and marginal revenue curves are shown by *DD* and *MR* respectively; the behaviour of average total costs is shown by *ATC*, and that of marginal costs by *MC*. If the firm desires the maximum net revenue per period that can be earned in the current conditions of demand for its product, it will plan to produce *OM* per period for sale at a price of *OP* per unit to earn a net revenue per period of *CBAP*.

Diagram 95 illustrates the sales plan that the individual firm will choose in the short-run, when operating in a market structure whose main characteristics are described by assumptions (*a*) to (*d*) inclusive on page 297 above. If existing firms are earning positive net revenues in the short-run, then in the long-run they may vary the quantities of plant, equipment, etc., that they

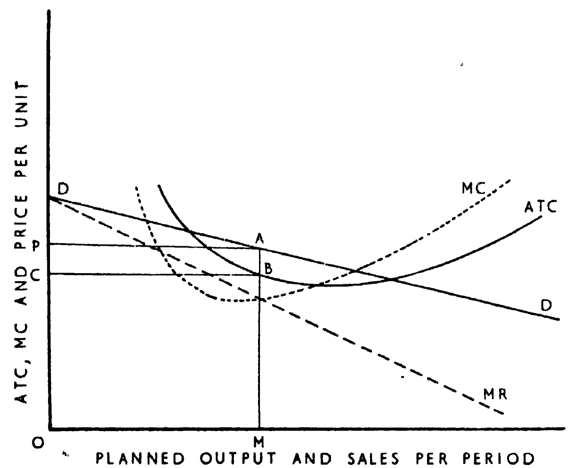


Diagram 95

use, and new entrepreneurs will plan to produce similar products. Each existing firm will be planning to earn the *maximum maximorum* rate of net revenue that can be extracted from the expected demand per period for its product in the long-run. It is not unreasonable to assume that the long-run demand for any firm's product will be more elastic at each price than the short-run demand for it.* If we ignore advertising expenditures, each firm will believe that the position of the long-run demand for its product is not affected by its present behaviour: since each firm acts independently of each other firm, and since the products that are currently being produced and those that may ultimately be produced by new firms are all equally substitutable for one another, no firm will have any incentive to behave like a monopolist who seeks to discourage potential competition, for by doing so it will forgo net revenue in the present and it will enjoy no compensating gain in net revenue in the future. The sales plan that each existing firm will hope to be implementing in each period in the long-run is illustrated in Diagram 96: *LD* and *LMR* are the long-run demand and marginal revenue curves respectively, and *LAC* and *LMC* are the long-run average cost and marginal cost curves.† The firm will be planning to produce *OM* per period at a price of *MP* per unit.

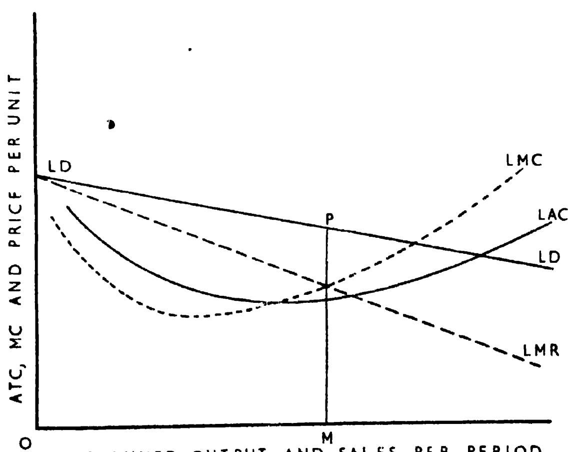


DIAGRAM 96

* See *supra*, pages 133–6.

† For a description of the derivation of these curves, see *supra*, Chapter 3.

As each new firm is set up in the long-run to produce a product that is similar to those being offered by existing firms, the demand curve for each existing product will shift negligibly to the left as some of its clients forsake it in favour of the new substitutes. The effect of continuous new entry will be continuous and appreciable falls in the demand for each existing product, and new entry will continue so long as the demand for existing products promise those who produce and sell them positive net revenues. It will only cease when each firm is implementing the sales plan illustrated in Diagram 97 — that is, when each firm is producing a rate of output of OM and selling it at a price of MP per unit, and in doing so, is just earning a revenue that covers its total costs of production in each period.

This simple model of monopolistic competition deviates from pure and perfect competition in two respects only: first, the product of any seller is not a perfect substitute for that of each other seller, and second, in explanation of product differentiation, no seller may produce a product that is a perfect substitute for that of any other seller. Thus far, then, differentiation of the products has been based 'upon certain characteristics of the product itself, such as exclusive patented features; trade-marks; trade names; peculiarities of the package or container, if any; or singularity in quality, design, colour, or style'.* Differentiation of the products may also arise because the productive ser-

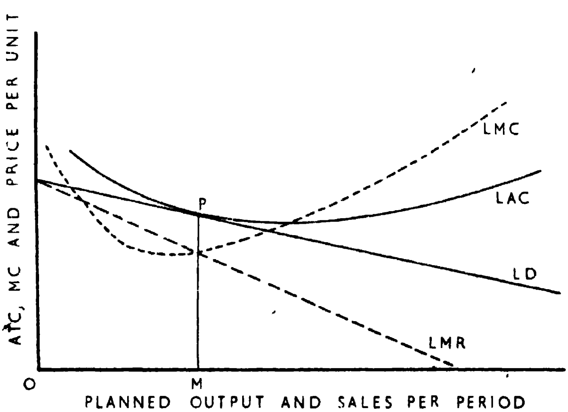


DIAGRAM 97

vices that any firm uses are not perfect substitutes for those being used by any other firm: thus, in retail trade, there may be differences between one firm and another in 'the convenience of the seller's location, the general tone or character of his establishment, his way of doing business, his reputation for fair dealing, courtesy, efficiency, and all the personal links which attach his customers either to himself or to those employed by him. In so far as these and other intangible factors vary from seller to seller, the "product" in each case is different, for buyers take them into account, more or less, and may be regarded as purchasing them along with the commodity itself.'* If the differentiation arises for these reasons, our analysis requires little modification: if the heterogeneous productive services are perfectly mobile between firms, then in the long-run each firm will be earning a zero net revenue as in Diagram 97; if it is the entrepreneurial factor that is heterogeneous, then in the long-run only the entrepreneur that is least 'efficient' will be in this position, and all the others will be earning positive net revenues which are commensurate with their relative efficiencies in the production and sale of this class of commodity.

There is no difference between the market behaviour of the individual firm that is operating under conditions of monopolistic competition and that of the monopolist in the simple model described on pages 281–5 above. Like him, the monopolistic competitor is a price-maker, and like a monopolist he may modify the sales environment in which he operates by introducing new variants of his product or by advertising. The monopolistic competitor differs from the monopolist, however, in that the sales environment of the former is being continuously circumscribed by the activities of all the other firms that are producing the same class of commodity, whereas that of the monopolist is (or can be made) tolerably secure. The sales environment of the monopolistic competitor is restricted by the existence of many close substitutes for his product: this means that the demand curve for his product will be more elastic at each price, and his range of choice in fixing its price consequently less, than that of the monopolist. In the long-run, his sales environment will be progressively restricted, for the demand for his product will decline and probably become more elastic as new

* Chamberlin, *op. cit.*, page 56.

substitutes for it appear, and these consequences of new entry can only be partly offset by advertising or by other efforts to bend consumers to his will.

Monopolistic competition, like monopoly and oligopoly, denotes a particular class of market structure. Whether or not any particular group of actual firms shall be filed in the box labelled 'monopolistic competition' depends on whether or not the market structure in which they operate is accurately described by the assumptions listed on page 297 above. The second of these assumptions — namely, that the product of any seller is an equally close substitute for that of any other seller — is the crucial one, for we must decide what is meant by 'equally close substitute' before we can check the empirical validity of assumptions (*a*), (*c*), (*d*) and (*e*): we must define 'close substitutes', for example, before we can count the number of firms that sell them or the number of households that buy them. In an exchange economy, all products compete with one another for the consumption expenditure of households; the degree of competition between them, however, would not merit our calling them close substitutes for one another. On page 297, we offered a temporary definition in terms of linguistic usage — namely, that products were close substitutes for one another if they were called by the same class-name, such as motor-cars or toothpaste. This, however, is unsatisfactory: a Rolls Royce and a Fiat 600 are both motor-cars, and while for many purposes they may be close technological substitutes for one another (in that they both provide transport), they are not close economic substitutes. This suggests that there must be some similarity in the prices of the products that are technological substitutes if they are to be regarded as economic substitutes. Alternatively, we may define equally close substitutes operationally — that is, in terms of the consequences that follow from their existence: in monopolistic competition, we want the different products to be such that, given the prices of $(n - 1)$ of them, the demand curve for the nth will be highly elastic, and we want it to shift appreciably to the left if there is a reduction in the prices of the $(n - 1)$ other products, and vice versa. That is, for each product we want a 'high' price elasticity of demand and a 'high' and positive cross-elasticity of demand between it and each other product in the same class.

This, however, poses a further problem: what cardinal num-

ber shall we take as typical of a 'high' price- or cross-elasticity? This question may answer itself if we find that there is a group of products within which cross-elasticities are markedly higher than they are between any product within the group and any other product outside it. If there is no such 'gap' in the chain of substitutes, then some critical value for the elasticities must be chosen. While a definition of close substitutes in terms of the price- and cross-elasticities is formally attractive, it may be of little use to the economist who seeks to catalogue market structures, for the data on which such calculations must be based will almost certainly be lacking. He must then tentatively delineate the products that he suspects to be closely substitutable for one another, using similarity in their prices and in the wants for which they cater as guides, and proceed to confirm his suspicions by uncovering the attitudes and expectations of the firms that are producing and selling them. Indeed, it is only by examining each case on its merits that an economist can decide whether the individual products are not only close, but also equal, substitutes for one another.

This requirement that each product must compete more or less equally with each other product within the group is the precondition of monopolistic competition that is probably least often fulfilled. Monopolistic competition may superficially seem to exist between shops that retail groceries within a large geographic area, for there is generally a large number of them, offering various standards of service; they are patronised by a relatively large number of buyers, and there is substantial freedom of entry. The geographic distribution of firms and households, however, is seldom such that all, or a sufficiently large number of households, can patronise with more or less equal convenience all or most of the shops. We usually find that the degree of competition between shops varies inversely with the distance between them: shop B, for example, may find that buyers regard its services as more or less equal substitutes for those of A, and of C; C may feel itself to be in close competition with B and D; D with C and E, and so on. In these circumstances, within the groups ABC, BCD, CDE, etc., each firm will be substantially affected by changes in the policies of the other firms within the group, so that we must classify each such group not as monopolistic competition but rather as oligopoly. If this

should happen — and it is probably not uncommon in the retail distribution of groceries, fruit and vegetables, petrol and pharmaceutical products — then instead of a single monopolistically competitive market we have a relatively large number of interlinked and overlapping oligopolistic markets.

We must not be dogmatic, however, about this distinction between monopolistic competition and 'chain' oligopoly. If we are concerned with the general consequences of some event for the grocery trade in general, it may be useful to argue as if all the grocery shops were operating under conditions of monopolistic competition. Thus, if we wish to predict the major consequences of the introduction or abolition of resale price maintenance, we will obtain predictions that prove to be in rough accord with what actually happens (or will happen) if we assume the patterns of behaviour that are characteristic of monopolistic competition. If, on the other hand, we wish to predict the possible results of a reduction of 2d. per gallon in the price of premium petrol by a particular filling-station in a particular district, it will generally be more useful to assume that some variant of oligopolistic competition obtains.

MONOPSONY AND MONOPSONISTIC COMPETITION

We shall deal briefly with these market structures, for they add little to the substance (though they add much to the variety of illustrative geometry) of our treatment of monopoly and monopolistic competition. For our 'ideal type' of monopsony, we shall suppose (*a*) that there is one buyer of productive service X; (*b*) that X is the only variable productive service that he buys; (*c*) that X is supplied to him by a purely competitive industry; (*d*) that he is only one of a large number of knowledgeable sellers of the product that X helps to produce, and (*e*) that the monopsonist believes that there is no possibility, either now or in the forseeable future, of other firms deciding to buy X. The relationship between the planned sales of X to the monopsonist and the price that he offers for it is the market supply curve of X. Since the monopsonist is the sole buyer of X, the supply curve of X describes the whole range of purchase plans that is open to him in each period: if he plans to buy relatively much of X he must offer a relatively high price per unit, and vice versa. The quantity that he decides to buy, and therefore the unit price that

he must pay for it, will depend on his objective. This choice is illustrated in Diagram 98, where we assume that the monopsonist seeks the maximum net revenue per period from his operations, that he knows the market supply curve of X, and the relationship between inputs of X and outputs of his product. The MM-curve shows the relationship between inputs of X and the marginal revenue productivity of X, and it is calculated in the manner described in Chapter 5. There, however, the MM-curve was the firm's demand curve for X for we assumed that the price of X lay beyond the firm's control; here, the MM-curve shows simply the additions to the firm's total revenue that would result from the use of successive units of X. The SS-curve is the market supply curve of X. The MSP-curve shows the amounts by which the firm's total expenditure on X would rise as successive units of X are purchased: thus, to buy OA units of X would cost the firm $OABC$, and to buy OD (= OA plus one unit) would cost it $ODEF$; the difference between $ODEF$ and $OABC$ (= the shaded area $ADEFCB$) is the amount by which the firm's total costs will rise as a result of its buying the OD-th unit of X, and this is represented by the distance DG.* The

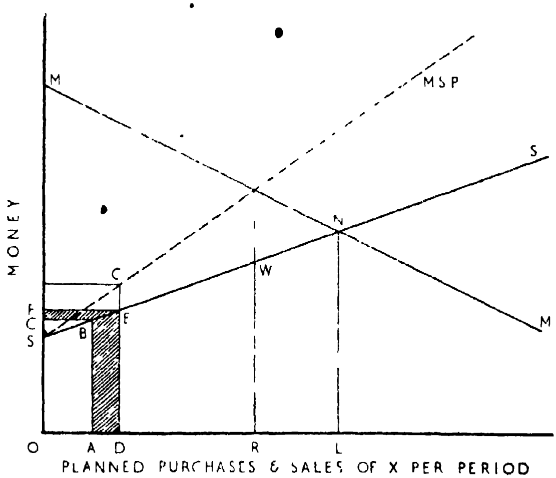


DIAGRAM 98

* The MSP-curve bears the same relationship to the SS-curve as does the marginal revenue curve to the demand curve. The MSP-curve must not be confused with the firm's marginal cost curve. The behaviour of the latter depends not only on the former, but also (in our simple model) on the relationship between inputs of X and outputs of the product.

monopsonist will earn the maximum net revenue per period by purchasing OR units of X at a price of RW per unit: if he were to purchase one unit more than OR his expenditure on X would rise by more than the additional revenue that he would earn by selling the products it helped to produce; to purchase one unit less of X would take more from his revenue than it would from his costs.

The simple model of monopsony may be modified in the same way as we modified the simple model of monopoly on pages 285–94 above. Thus, the monopsonist is not obliged by the market structure to pursue the maximum net revenue per period: irrespective of the objective he chooses, however, his planned purchases of X per period are unlikely to fall below OR or rise above OL, and its price is unlikely to fall below RW or rise above LN. If X is being produced in several, independent, purely competitive markets, and if the elasticities of supply of X vary between them, the monopsonist may enhance his net revenue by paying different prices in different markets, and the price he pays will be lower in the market where X is in relatively elastic supply, and higher in the market where X is in relatively inelastic supply. It will pay the monopsonist to divide the market in which he buys X into sectors between which no transfers of X are possible, provided the costs are less than the additional revenue he expects to earn from doing so. Lastly, it may pay the monopsonist to advertise for new sources of X and so shift the market supply curve of X to the right.

We have seen (*supra*, page 297) that monopolistic competition may exist if buyers are not indifferent as to which seller they patronise; if, *inter alia*, sellers are not indifferent as to which buyer they sell to, we may have monopsonistic competition. For our ideal type of monopsonistic competition, we shall suppose (*a*) that there is a very large number of sellers of a productive service, S, which may be a particular kind of labour-service; (*b*) that each unit of S is a perfect technological substitute for each other unit; (*c*) that there is a large number of firms buying this productive service and that it is the only variable productive service that they buy; (*d*) that each firm is a price-taker for the product it sells and for each other input that it buys; (*e*) that sellers are not indifferent as to which buyer they offer their services, and (*f*) that the only limitation on the entry of new buyers

is that there may appear no buyer who is identical in the estimation of sellers with any existing buyer. Given these assumptions, the supply curve of *S* to each buyer will be highly, but not perfectly, elastic: if he offers a higher price, more (but not all) sellers will patronise him; if he offers a lower price, only some of those who now supply him will forsake him. The choice of a purchase plan by an individual buyer is illustrated in Diagram 99(*a*), where we assume that he knows his demand for *S* and the supply of *S* to him. The sales plan for its product is shown in Diagram 99(*b*). The firm will be buying *OR* of *S* at a price of *RW* per unit in each period and producing with this (in conjunction with the fixed quantities of other services at its disposal) an output of *OQ*

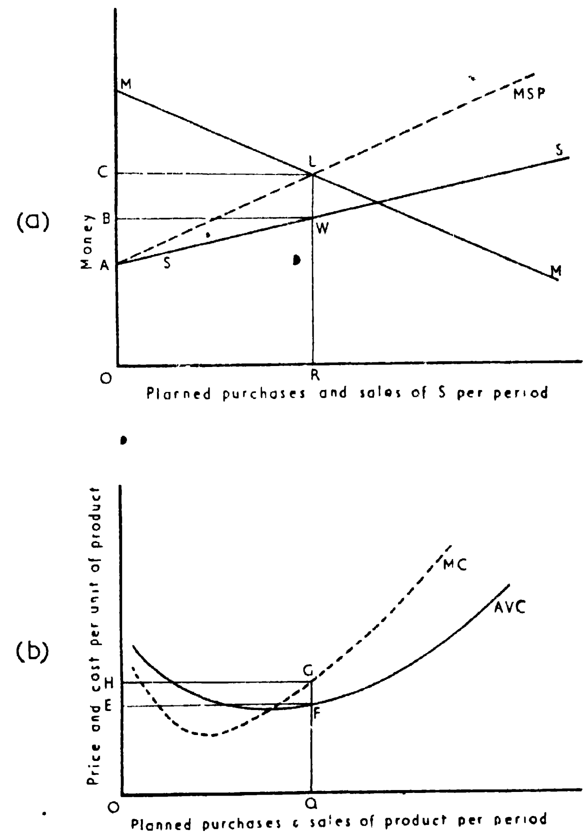


DIAGRAM 99

per period; when doing so, it will be earning the maximum net revenue per period.* The excess of total revenue over total variable costs is shown by the areas ALM and $BWLC$ in Diagram 99(a) and $EFGH$ in Diagram 99(b). If this excess is more than enough to cover the fixed costs of existing firms, and if any new firm can enjoy all the advantages that are being enjoyed by existing firms (save that of being equally esteemed by the sellers of S), then new sellers of the product (buyers of S) will appear in the long-run. As new entry proceeds, the price of the product will fall, and as this occurs, each firm's demand for S will shift to the left; furthermore, the supply of S to each firm may shift leftwards and become more elastic. These adjustments will continue until each firm is earning an excess of total revenue over total variable costs in each period that just suffices to cover its fixed costs — that is until each firm is earning a zero net revenue.

A NOTE OF PROFITS*

By the 'profits' of a firm we mean (roughly) the actual net revenue that it has earned in some past period of time. We have not used this term thus far, and for two reasons. First, in the previous chapters we have been concerned more with the choice of plans than with the results of their implementation. Second, the market morphology that was implicit in Chapters 1–8 was (generally) one of pure and perfect competition, and when these conditions obtain profits tend towards zero as long-run adjustments are effected. There is a presumption, therefore, that positive profits are a consequence of 'imperfect'† competition, and this explains why we have postponed until now any attempts to define profits and list the things on which their magnitude depends.

We shall distinguish between 'profit' and 'profits', and we shall define both terms operationally — that is, in such a way that they will do the job that we want them to do in our analysis.

By profit we mean the net revenue that a firm expects to earn during a period of time that lies ahead; by profits we mean the net revenue which a firm has actually succeeded in earning during a period that has ended. We have assumed that the firm decides to implement the sales and purchase plans that promise it the maximum profit. If the expectations on which these plans are based — that is, the expected selling prices of the products and the expected prices and productivities of the productive services, etc.,— are fulfilled as time passes, then when the period has ended the firm will have earned the maximum profits. A comparison between profit and profits in a particular period, then, provides a measure of the extent to which the firm erred in the estimates on which its plans were based; if the economic environment is relatively stable over time we would expect the difference between profit and profits to dwindle and disappear.

In economic analysis, we want the expectation of profit and/or the earning of profits by the firms already established in an industry to cause new firms to be set up to compete with them in the long-run, and we want the expectation of negative profit or the presence of negative profits (that is, losses) to cause the withdrawal of some firms from that industry* in the long-run. If profit is to allocate entrepreneurs (and therefore other productive resources) amongst different industries, we must measure it in such a way that the profit which an existing firm expects to earn is identical with that which a firm that was quite new to that industry might expect to earn were it set up now and were it to use the same methods of production and to enjoy the same advantages as the firm that is already there. This means that a going firm, when estimating its profit, must deduct from the total revenue that it expects to earn the costs of all the productive services that will be used to produce the products whose sale will yield that revenue, and its calculations of the costs of these services must be based on their current market prices. That is, it must calculate its costs for the ensuing period on the same basis

* The only unambiguous definition of an industry is a group of firms each of which is producing a product that is a perfect substitute for that of each other firm. If the 'product' is not perfectly homogeneous, precise delineation of the limits of the group is impossible. We have alluded to this problem earlier on pages 302–3. For our present purposes, it is sufficient to define an industry as a group of firms producing a substantially similar product(s) with the help of substantially similar productive resources.

as it would use if it were now replacing all the resources that are now at its disposal. This means that the services of the entrepreneur for the ensuing period must be valued at the minimum sum of money that would be needed to retain him where he now is, and this will be equal to the maximum earnings that he could obtain were he now quite unencumbered by his past decisions and therefore free to choose what industry to enter. We know (*supra*, Chapter 3) that the cost of the firm's 'fixed' factors in each period is made up of (*a*) the payments that have to be made in each period by way of interest and dividends on the monies that were used to buy them, and (*b*) the depreciation, which in its turn depends on the expected life of the fixed factors, their prices and the rate of interest that the firm can earn on the sums of money set aside for depreciation while they are being accumulated to finance replacement. When estimating its profit for the ensuing period, the firm must base (*a*) on the rate of interest that it would now have to pay to induce people to buy its bonds, and on the rate of dividend that would now have to be promised to induce the public to buy its ordinary shares; it must base (*b*) on the replacement costs (that is, the current market prices) of its fixed factors and on the maximum rate of interest that it could earn currently on monies set aside by way of depreciation. If an existing firm estimates its profit in this way, then its profit will be identical with that which a new entrepreneur, who was equally versatile and efficient as the one that was already established in the industry, would earn were he to enter that industry now, using precisely the same quantities and qualities of productive services.

When profit is calculated in this way, and when an industry is perfectly competitive,* the profits actually earned by each firm will be zero in each period in the long-run. If there are no institutional or legal barriers in the way of new firms entering the industry, and if all productive services are perfectly mobile and divisible, and if their owners have both perfect knowledge and foresight, then the profit estimate of each potential entrant into the industry will be identical with that of each firm that is already established in it, and the influx of new firms (each in all

* In the rather extreme sense in which we defined perfect competition in the previous chapter — namely, that an industry is perfectly competitive when all the productive services that it uses (including entrepreneurs) are in perfectly elastic long-run supply to it.

respects the same as each existing firm) will continue until no owner of any productive service can earn a higher reward there than elsewhere. This result is not so much a consequence of the extreme values that we have given to the variables: legal barriers, mobility, divisibility, knowledge and foresight, when defining perfect competition, but rather of the fact that each variable has the same value for each owner of any productive service, for this suffices to ensure that no existing firm will possess any advantage as compared with any potential new entrant.

If firms inside an industry enjoy no relative advantages as compared with outsiders that contemplate entering it, then the profit estimates of outsiders and insiders will be identical, and will tend towards zero in the long-run. If the firms that are already established in an industry enjoy relative advantages as compared with potential entrants, their profit estimates will exceed those of outsiders, and in the long-run the established firms will be earning positive profits, for new entry will only continue until the maximum profit that any outsider expects to earn is zero. The explanation of profits, then, must lie in the relative advantages that the firms that earn them enjoy as compared with other firms that might be set up to compete with them. These advantages may assume many forms. Established firms may be protected from new entry by institutional or legal barriers; their entrepreneurs may be more knowledgeable or they may be less deterred by uncertainty; they may be able to buy a necessary productive service at a price which is lower than what newcomers would have to pay because it is immobile, or their relative advantage may lie in indivisibilities of productive plant and processes. We shall now examine briefly how each of these might explain the existence of profits in the long-run.

First, obstacles in the way of new entry. If the entry of new firms into an industry were prohibited by statute, profit estimates of the firms within it may be high and positive indefinitely. The same may apply where new entry is regulated by licence and where the licences are not freely issued to all who ask for them: the profit estimates of outsiders, who lack licences, may be zero or negative while those of established firms are positive, so that the latter continue to earn profits. If the existing firms in an industry have patented their products or processes, new firms will be prevented from being set up to compete with

them, and the expectation and fact of profits may continue until the patents expire. If the prohibition, licence or patent is the sole deviation from conditions of perfect competition, we may call the resulting profits 'monopoly profits' or 'monopoly rents'. Our present purpose is to explain the magnitude of profits, and not to isolate the notional components — namely, economic rents and 'pure' profits — into which they may be divided. If we wish to do this latter, however, a useful criterion is whether or not the 'cause' of the profits can be transferred from one person (or firm) to another. In the examples quoted earlier, this is possible, for the right to operate in an industry to which entry is prohibited, or a licence or a patent required, can be bought and sold. If entry into the industry is otherwise free, the sale of any of these will yield a capital sum such that the interest per period on it at the current and relevant rate will just be equal to the profits that were being earned in each period by the seller. Since the right, licence or patent has no value per period in any other use, the profits (that is, the 'price' at which it is currently being 'sold' in its existing use per period) may be regarded as economic rent. If a 'cause' of profits cannot be bought and sold, we shall call that part of the profits that it explains 'pure' profits. We shall return later, however, to the notional components of profits.

Second, immobility of productive services. If some resource X is perfectly immobile, it will be in perfectly inelastic supply to its existing use. We shall distinguish two cases: first, where X is immobile because its owners will not (or can not) move to other occupations, and second, where X is in inelastic supply because the owners of other productive services that might be made substitutable with X in the long-run will not (or can not) change these services into a form in which they will augment the supply of X. The first case is illustrated in Diagram 100. We suppose for the sake of simplicity that there is only one firm buying the immobile resource and that it is a price-taker for its product(s); DAD' is the firm's demand curve for X, where OD is the maximum average revenue productivity of X to the firm and AD' is that part of its marginal revenue productivity curve for X that lies below the maximum point of its average revenue productivity curve,* and SS' is the supply curve of X. The firm is buy-

* See *supra*, Chapter 5.

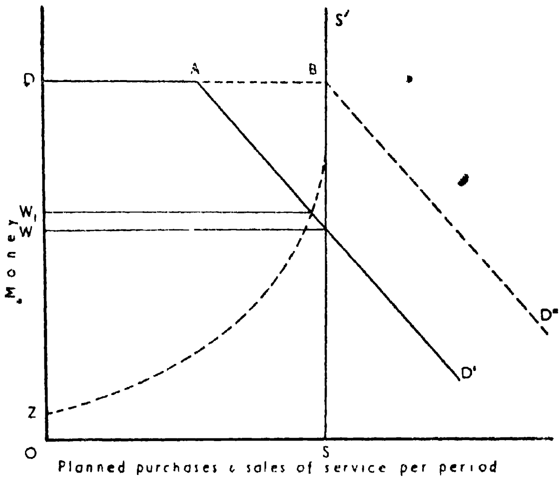


DIAGRAM 100

ing OS of X at a price of OW per unit, and since OW is less than the average revenue productivity of that quantity of X the firm will be earning profits in each period. In the short-run these profits are a consequence of the immobility of X. If some owners of X were to move to other occupations in the long-run, so that its long-run supply curve became ZS', its price would rise to OW', and the firm's profits per period would fall. If there were perfect competition in the long-run, the total demand curve for X would rise to DBD'', and the price of X would rise to OD; at this level, the profits of each firm would be zero for the price paid per unit of X would be equal to its average revenue productivity. If the profits persist in the long-run, therefore, the immediate cause must lie either in the continued immobility of X or in the absence of free entry; of these the latter is paramount for it alone can explain the presence of profits. The fact that new firms are not set up to compete with that which is now buying X may be due to the presence of institutional or legal barriers, lack of knowledge, indivisibility or uncertainty. We have already seen that profits persisting for the first of these reasons are in the nature of economic rent, and we shall presently explore the notional components of profits that arise for any of the other reasons. Immobility in this first sense, then, will not explain persistent profits, for their cause must lie in the reasons why new

firms do not appear to bid up the price of the immobile resource to the level at which the firms that use it do not earn profits.

The effects of immobility in our second sense are illustrated in Diagram 101. We suppose that SS' is the short-run supply curve of the service of a 'fixed' factor, such as a piece of equipment, that is owned by an existing firm. The sum of money OZ represents the 'fixed' costs of these services — that is, the 'price' of the equipment's services calculated for each period on the basis of the current rates of interest or dividend and the current market price of the equipment. If we ignore indivisibilities, the long-run supply curve of these services will be ZZ'. The DD-curve is the firm's demand for these services per period, and it is derived from the demand for the product(s) that they assist in producing. The current value of the equipment's services to the firm will be OP, and it will be earning profits in each period, for its profits are calculated not on the basis of the current marginal revenue productivity of the equipment but of its replacement cost. If entry into the industry is perfectly free, the marginal revenue productivity of the services of each unit of equipment will fall until it is equal to its 'fixed' costs in each period, and no firm will be earning profits. With free entry, therefore, such profits are temporary. If an existing firm were being sold as a going concern, the price that would be paid for it would reflect

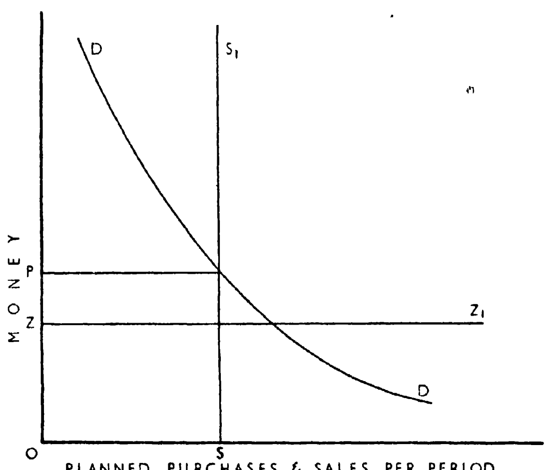


DIAGRAM 101

the existence of these profits, and for this reason we may regard them as economic rent or quasi-rent.* If profits that arise from this kind of immobility are to persist, then the entry of new firms must be prevented. Immobility in this second sense, therefore, is a mediate, rather than an ultimate, cause of profits.

Third, the profit estimates of entrepreneurs outside an industry may be zero or negative, while those of the firms that are now operating within it may be positive, because the insiders are more knowledgeable than the outsiders. This divergence will be temporary if the knowledge (for example, of plant lay-out or productive processes) that going firms possess, but which potential entrants lack, is freely accessible to all. The divergence can only be permanent if the relevant knowledge is permanently denied to outsiders, and this can be done by patents (and other devices) that restrict knowledge, or the right to apply knowledge, to existing firms. In this case, the profits are explained by the patents that prevent the free flow of knowledge.

Fourth, indivisibilities in organisation, machines, equipment or processes may explain why established firms earn profits in the long-run. We shall deal with indivisibility briefly, for we have already (*supra*, pages 270–2) alluded to its probable effects. Let us suppose that there are five firms in an industry, that each is using the quantities and kinds of productive services that promise it the lowest average costs per unit of its output, and that the demand for their product is such that each is earning zero profits. Let us now suppose that the planned purchases of the product by households rises by 5 per cent at each price. As a consequence, each existing firm will expect to earn a positive profit. The profit estimate of a potential entrant to the industry, however, may be negative: if he were to enter the industry on the same scale as the firms that are already there, the increase in the total output of the product would be of the order of 20 per cent, and the price at which the output (so enhanced) could be sold might be so low that both the new firm and the established firms would earn negative profits. An outsider will only expect a zero profit or a positive profit if the demand for the product has risen by the same (or a larger) proportion as that by which his entry would increase the output of the product. The profits that arise from indivisibility may be called 'monopoly' profits, where

* Where 'quasi' draws attention to the fact that they are temporary.

by 'monopoly' we mean the protection of the markets of established firms from the encroachments of new entrants. These profits may also be regarded as economic rent: the fact that a firm is already established in an industry where indivisibilities are important is as much one of its assets as 'good-will' or the possession of a licence, and if the firm were sold we would expect the price that was paid for it to include the present value of the stream of profits that was expected to accrue to the firm in the future as a result of its ownership of the indivisible resource. These profits are of the same kind as those which the firms that are already established in a perfectly competitive industry may earn in the short-run; they differ in degree in that these short-run profits are temporary while those arising from indivisibility may be permanent.*

Fifth, the profit estimates of entrepreneurs who are outside an industry may be zero or negative, while those of established firms are positive, because insiders and outsiders are not equally uncertain about what the future holds in that industry, or because, being equally uncertain, they have different attitudes towards uncertainty. We have already seen how uncertainty and the attitudes of the owners of any productive service towards it may by themselves explain why that service is in less than perfectly elastic supply in the long-run. Here, we shall concentrate on the way in which uncertainty may explain the divergence between the profit estimates of established firms and of potential entrants to an industry, and so explain why established firms earn profits in the long-run. The profit estimate of an entrepreneur is the difference between his expected revenue and expected costs† per period: the former depends on the expected behaviour of the demand for the product that he is producing (or hopes to produce); the latter depends, *inter alia*, on the prices he must pay for his 'fixed' factors, the prices he expects to have to pay for the variable productive services that he must employ, and on his estimate of their physical productivities. Lacking perfect foresight, he cannot know what value(s) each of these will assume as he puts his plans into effect. The choice of what

* By permanent we mean that they will continue for as long as it takes the demand for the product to rise by a proportion large enough to justify the entry of a new firm.

† Costs as defined on pages 309–10 above.

sales and purchase plans to implement, however, can only be made when some value is given to each of them.

We know little or nothing about the actual mental processes by which a businessman, who cannot know what the future holds, decides upon his present plans. As economists, we can only observe his actual behaviour in the face of uncertainty, and it is probable that this is what it would have been had he attached a 'best guess' value to each determinant of his future revenue and costs, and had he discounted each of these to a present and 'certain' value at a discount rate which reflected his attitude towards uncertainty. Thus, to say that a businessman cannot know the price at which he will be able to sell his product *n* periods hence is to say that the selling price of the product may assume any one of the many values $a_1, a_2, a_3, \ldots a_r$. He cannot attach a cardinal number to the probability of any one of these actually occurring, but he may feel that some values are more (or less) likely than others. We shall assume that he in some way reduces this range of likely values to a single and 'most likely' value, $\bar{a}$; that he deducts from this a safety-margin, c, to guard against the possibility that his best guess will not be fulfilled, and so obtains a 'certainty-equivalent' $(\bar{a} - c)$. We shall assume that he follows the same procedure in calculating his costs, adding a safety-margin to his 'best guess' costs, and that he estimates his profit by deducting the 'certainty equivalent' of costs from that of revenue. Profit estimates that are made in this way will vary from one entrepreneur to another: they may vary because (*a*) the range of possible values for each element of revenue and costs may be different for different entrepreneurs; (*b*) even if the ranges are identical, there will be variations in the 'best guess' values to which they are reduced, and (*c*) even if both the ranges and the best guesses are the same, different entrepreneurs may use different safety-margins. If entrepreneurs who contemplate entering an industry have a low range of possible values for revenue and a high range of possible values for costs, if they have a low 'best guess' value for revenue and a high 'best guess' value for costs, and if they use large safety-margins, then their estimate of the profits they might earn were they to enter that industry will be zero or negative. In these circumstances, no new entry will occur, and if the firms already established in the industry are both earning profits and expecting profit, they

may, *ceteris paribus*, continue to do so. The profits of existing firms are explained here by the greater caution of potential entrants: if there had been no uncertainty, or if outsiders and insiders had viewed the future with 'equal' uncertainty and with the same attitudes towards uncertainty, then the profits of established firms would have dwindled and disappeared as new entry took place.

Profits that arise for this reason are 'pure' profits and not economic rent. They are earned by a firm because its entrepreneur makes better 'best guesses'* and calculates with lower safety-margins than other entrepreneurs. They are a reward to him for operating in an industry from which others are repelled by uncertainty. If a firm which had been earning these profits were sold, their magnitude would not be reflected in the price that was paid for it, for the purchase price would depend on the buyer's assessment of the uncertain future. In this respect, they are unlike the profits that arise from indivisibilities or from licences and patents, and for this reason we call them 'pure' profits, where 'pure' means that they are not economic rent. We may speak of these profits as being the entrepreneur's reward for 'bearing' uncertainty: the reward is positive (or higher than it otherwise would have been) because others are unwilling to bear this uncertainty, in the same sense as the wage-rate at which a carpenter can sell his services is higher than it would have been had more workers been willing to acquire the skill of carpentry. These profits are not a measure of the uncertainty that is borne by the entrepreneur that earns them. If they measure anything, they measure the uncertainty as it is seen by the entrepreneurs who refuse to 'bear' it — that is, as seen by those who are deterred by it from entering the industry. These profits will accrue to an entrepreneur for as long as he continues in the industry: uncertainty and his attitude towards it bear no relation to the size of his profits, provided that the safety-margins which he uses are such as to promise a profit that will induce him to remain in the industry. If we want a measure of the uncertainty that an entrepreneur has borne, we may find it in the difference between his profit and his profits for a particular period of time, for the difference between what he expected to

* That is, best guesses that are more often proved to have been correct by events.

earn and what he in fact did earn in that period will reflect the extent to which what he expected to happen did not happen.

We may summarise the conclusions of this section briefly. When an industry is perfectly competitive — that is, when entry into it is perfectly free — each firm in it will be earning zero profits in the long-run. Indeed, an economist when contemplating such an industry may well use the words of the Preacher, the son of David: 'Then I looked on all the works that my hands had wrought, and on the labour that I had laboured to do: and, behold, all was vanity and vexation of spirit, and there was no profit under the sun.'* Profits are absent in the state of perfect competition because each entrepreneur has equal opportunities and equal capacities to exploit them. The presence of profits in the long-run, therefore, means that different entrepreneurs are aware of different opportunities and that their abilities to exploit the opportunities that are open to them are different also. These differences may arise because (*a*) entrepreneurs already established in an industry may be protected by law or by (other) barriers that they themselves have created; (*b*) entrepreneurs already there may be 'protected' by 'indivisibilities' of plant and processes, and (*c*) entrepreneurs do not feel equally uncertain about the future and are not equally willing to live in the presence of uncertainty. These, then, are the 'causes' of profits in the long-run. The profits that are explained by (*a*) and (*b*) are of the nature of economic rent; those explained by (*c*) are not economic rent and for that reason were called 'pure' profits. Finally, we should note that the strength with which any one of these causes operates may not be independent of the magnitude of each of the others: thus, the presence of legal and institutional barriers in the way of new entry may diminish uncertainty for the firms that are protected by them, and the presence of 'indivisibilities' in an industry may increase uncertainty for those who contemplate entering it.

* Ecclesiastes, Chapter II, Verse 11. (Holy Bible, Authorised Version.)

CHAPTER 11

Oligopoly and Bilateral Monopoly

OLIGOPOLY

All markets in which there are a small number of sellers are classified under the heading 'oligopoly'. The adjective 'small' must be interpreted operationally: the number of sellers of a homogeneous or differentiated product must be such that each believes that any change in his selling price and sales, or in the quality of his product, or in his advertising expenditure, or in any other variable whose value is under his control, is likely to evoke retaliation from most or all of the other sellers. When the number of sellers is small in this operational sense, we generally find that there is a small cardinal number of sellers — that is, not less than two and perhaps not more than twenty. It is for this reason that economists decide whether or not to classify any particular market as an oligopoly by counting the number of firms: this provides a recognisable, objective and measurable criterion for classification, whereas the awareness of mutual interdependence of sales, purchase, production and advertising plans is less easily established, for it is always a matter of degree and frequently a matter of opinion.

In this section, we shall study a number of models of oligopolistic markets. Each model explores the probable consequences of a particular assumption that is made by each oligopolist about his rivals' reactions. The models are not listed in any simple logical order. The order in which they appear below is roughly one of increasing knowledge by each oligopolist about his rivals' reactions. Since each oligopolist is the more likely to make a correct assumption about rivals' behaviour — that is, an assumption that accurately describes their reactions as he acts on the basis of it — the greater is the degree of tacit or overt agreement between them, the order is also roughly one of increasing degrees of agreement or collusion. The greater the degree of collusion, however, the nearer might oligopoly approximate towards

simple monopoly, so that the order in which we list the models is very roughly one of increasing profits for some or all of the firms.

THE COURNOT MODEL*

In this first model, we shall suppose that (*a*) there are only two independent† firms — that is, there exists the simplest example of oligopoly, namely, duopoly; (*b*) each produces and sells a product that is a perfect substitute for that of the other; (*c*) the product is perishable and cannot be stored, so that in each period the total output of it must all be sold; (*d*) there are many knowledgeable buyers of the product; (*e*) each duopolist knows the market demand curve for the product; (*f*) the two firms have identical cost curves, and to simplify the geometry we shall assume that for each duopolist the cost of production for each output is zero; (*g*) each duopolist makes an output plan at the beginning of each period setting out the quantity of the product which he plans to produce during it, and once made, an output plan cannot be revised; (*h*) neither sets a price for his output, but each accepts the price at which the total planned output can be sold, and (*i*) each duopolist seeks the maximum net revenue in each period. Lastly, we shall suppose that (*j*) while the duopolists are aware of the mutual interdependence of their output plans, each is quite ignorant of the direction and magnitude of the revision in his rival's plan that would be induced by any given change in his own; each, however, in making his own plan must make some assumption about his rival's reactions, and we shall suppose that each duopolist assumes that irrespective of the output plan that he implements in any period $t+1$, his rival will maintain his output at the same level as in period t.

Given assumptions (*a*) to (*h*) inclusive, we can illustrate the range of profit possibilities that is open to each firm. In Diagram 102, the *DD*-curve shows the total quantity of the product that households would plan to buy in each period at each price at which it

* See Augustin Cournot, *Recherches sur les principes mathematiques de la theorie des richesses*, Paris, 1838. There is an English translation by N. T. Bacon, entitled: *Researches into the Mathematical Principles of the Theory of Wealth*, New York, Macmillan, 1897.

† By 'independent' we mean that there is no kind of agreement or collusion between them.

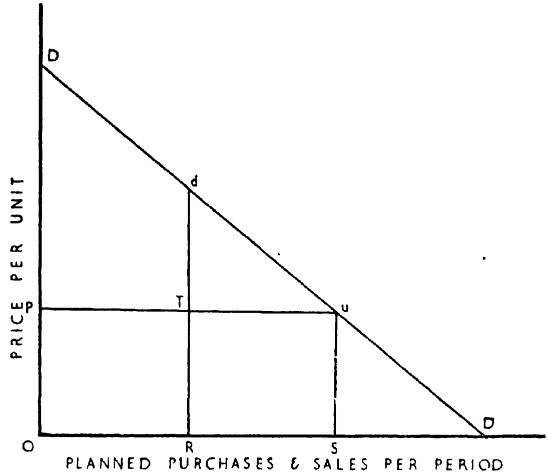
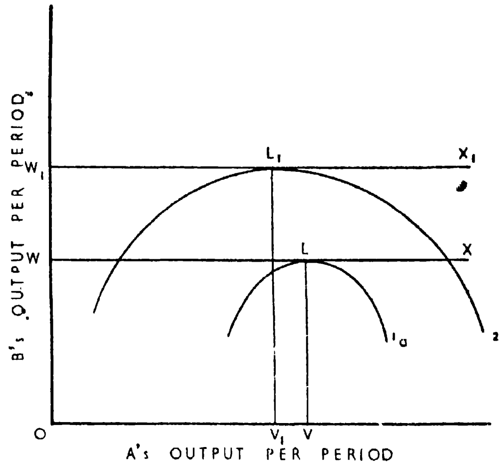


DIAGRAM 102

might be sold. In Diagram 103, we measure the alternative outputs that duopolist *A* might produce in each period on the horizontal axis, and on the vertical axis, we measure the output of his rival *B*. Any point that lies between these axes represents a particular combination of the output of *A* and of *B*. Beside each point we can write the profits that each duopolist would be earning when they are producing the outputs which that point denotes. Thus, the point *L* denotes an output of *OV* per period by *A* and of *OW* per period by *B*. By transferring these outputs to the horizontal axis in Diagram 102, we can discover the selling price per unit: thus, *OR* is equal to *OV* and *RS* to *OW*, and the total output (*OV* plus *OW*, or *OS*) can be sold at a price of *OP* per unit. Since we have assumed that each duopolist has zero costs of production, *A*'s profits are represented by the area *ORTP*, and *B*'s profits by the area *RSuT*. In precisely the same way, we may obtain the profits that each of the duopolists would be earning were they producing the outputs denoted by any other point lying between the axes in Diagram 103. When this has been done for each point, we obtain a visual representation of the profit possibilities open to *A* and to *B*. We can order the profit possibilities that are open to either duopolist by drawing profit-indifference or iso-profit curves, each of which passes through all combinations of *A*'s and *B*'s output which promise *A* (or *B*) the same sum of profits per period. In Diagram 104,



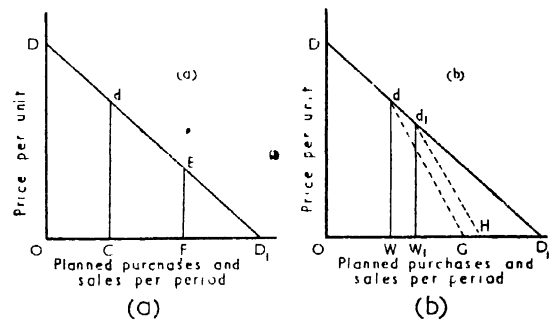


DIAGRAM 103

the profit-indifference curves of A and B are drawn and we may easily explain the shape that we have given them.

Let us take any value for B's output, and, keeping this constant, examine what happens to A's profits as A's output increases. In Diagram 103(a), DD_1 is the market demand curve for the product, and OC is B's output. The relationship between A's output and the selling price of the product is shown by the range dD_1 of the demand curve: since this curve is relatively elastic between d and E, A's total receipts (which are also his profits since his costs of production are zero) will rise as his output per period increases from zero to CF; at prices lower than FE, the dD_1-curve is relatively inelastic, so that as A increases

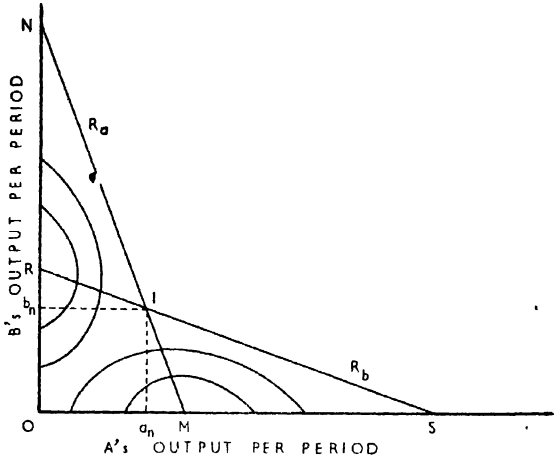


DIAGRAM 104

his output per period from CF, his profits will continuously decline, reaching zero when his output is CD_1. Next, let us take any value for A's output, and examine what happens to A's profits as the value of B's output rises. In Diagram 103(a), if we suppose that OC represents A's output, it is clear that A's profits will continuously decline as B increases his output from zero to CD_1, for A's profits are represented by the area bounded by OD, OC, Cd and a horizontal line drawn at the selling price of the product, and as B's output rises the price of the product continuously falls so that this area becomes progressively smaller. These two conditions — namely, that at any value for B's output, A's profits will rise, reach a maximum and then decline as A's output is increased, and that at any value for A's output, A's profits will continuously fall as B's output is increased — are both fulfilled by profit-indifference curves that are concave when viewed from the axis on which we measure A's output. Lastly, we must explain why the maximum points of successive iso-profit curves of duopolist A lie progressively nearer to the axis on which we measure B's output. If B is producing the output OW in Diagram 103, the alternative profits that A might earn by varying his output will lie on the line WX; if A seeks the maximum profits per period, he will plan to produce the output at which this line is tangential to one of his iso-profit curves. Let us suppose that when B is producing OW per period, A's profit-maximising output is OV — the output where the line WX just touches the

maximum point L of the profit-indifference curve $1a$. If B's output were higher at OW_1 per period, then the output at which A's profits would be greatest would be OV_1 — the output at which the line W_1X_1 touches the maximum point L_1 of the iso-profit curve $2a$.

The output OV_1 is less than OV, and we can quite easily confirm why this must be so from Diagram 103(*b*). When B is producing OW, A's profits will be at a maximum when A's output is $WG\,(=OV)$, the output where the marginal revenue curve corresponding to the range dD_1 of the market demand curve cuts A's marginal cost curve; when B's output is OW_1, A's profit-maximising output will be $W_1H\,(=OV_1)$, the output where the marginal revenue curve corresponding to the range d_1D_1 of the demand curve cuts the horizontal axis which is A's marginal cost curve. Since in our example the market demand curve is a straight line, $WG=\frac{1}{2}WD_1$, and $W_1H=\frac{1}{2}W_1D_1$; W_1D_1 is less than WD_1 so that $W_1H\,(=OV_1)$ must be less than $WG\,(=OV)$. The profit-maximising output of A will be lower, therefore, the higher is B's output per period: that is, the apices of A's iso-profit curves must lie progressively nearer the axis on which B's output is measured. The general properties of B's iso-profit curves may be established in a precisely similar fashion.

The assumption that A (or B) makes about B's (or A's) reactions when deciding what output to produce in the ensuing period can be illustrated in Diagram 104, wherein are drawn the profit-indifference maps of A and B. If A assumes (see assumption (*j*), *supra*, page 321) that B will always maintain his output at its level of the previous period irrespective of the output which he (A) produces, then the profit-maximising output of A for each level of B's output will lie on the line MN which passes through the maximum points* of A's iso-profit curves. The line MN is called A's reaction curve, for it shows us how A will react to any change in B's output. In Diagram 104, A's reaction curve is a straight line because we have assumed that the market demand curve is a straight line and that A's marginal costs of production are constant (at zero). The output OM is the 'monopoly' output, for it is that which A would plan to produce if B's output were zero — that is, if A were the sole producer and

* By the 'maximum point' of any one of A's iso-profit curves, we mean the point furthest from the horizontal axis on which A's output is measured.

seller of the product. The output ON is that which B would have to produce to induce A to choose a zero output. We can see from Diagram 103(a) that ON must be the output at which price and marginal cost are equal, and since this is the output that would be offered for sale in each period had the product been produced under conditions of pure competition — that is, by many firms each of which had zero costs of production — we shall call it the 'competitive' output. Since A and B are in all respects identical, OM will be equal to OR and ON will be equal to OS.

If each duopolist seeks the maximum profit per period and if each assumes that his rival's output will be maintained at its level of the previous period irrespective of the output which he now produces, then they will ultimately be producing the outputs denoted by the point I at which the two reaction curves intersect one another — that is, A and B will be producing Oa_n and Ob_n respectively. This necessarily follows from our assumptions (a) to (j) above. We can trace the path by which the equilibrium denoted by I is reached by supposing that A is initially a monopolist and that a competitor B suddenly and unexpectedly appears at the beginning of period 1. In period 0, A's output is OM in Diagram 105, where the reaction curves appear uncluttered by the profit-indifference maps. The new

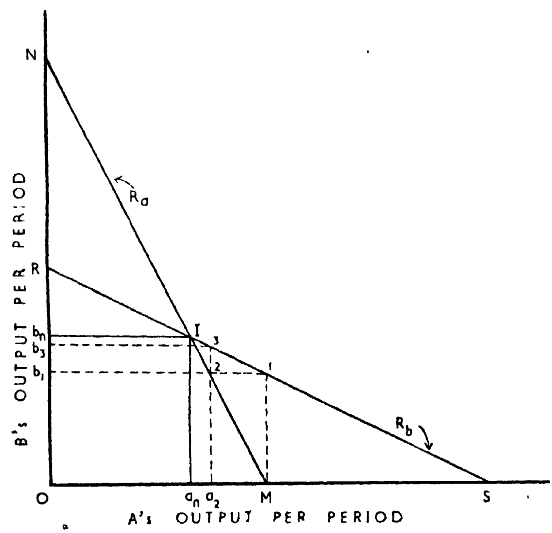


DIAGRAM 105

firm, B, assuming that A will continue to produce OM in period 1, plans to produce Ob_1; A, assuming that he will have no rival, plans to produce OM. The combination of outputs that will actually be produced in period 1 is denoted by the point 1 on B's reaction curve. In period 2, B will plan to produce Ob_1, for that is the output that promises him the maximum profits if A produces OM, and he assumes that A *will* produce OM; A will plan to produce Oa_2 since he assumes that B will maintain his output at Ob_1. In period 2, the total output of the product will be Oa_2 plus Ob_1 — that is, that denoted by the point 2 on A's reaction curve. In period 3, if each duopolist continues to take the output of his rival in the previous period as a datum, A and B will produce Oa_2 and Ob_3 respectively. It is clear from the diagram that these adjustments will continue until A and B are producing Oa_n and Ob_n respectively.

The Cournot model is analytically attractive because it yields a unique and stable equilibrium for each duopolist. This equilibrium is denoted by the point I in Diagram 105 where the two reaction curves intersect one another. The nature of this equilibrium is largely determined by assumption (j) and its precise content is mainly explained by assumptions (a) to (i). If we maintain the former assumption and if we vary the latter assumptions within the general framework of an oligopolistic market, we still get a single equilibrium, provided that the output which each oligopolist would produce if each of his rivals was producing nothing is less than the outputs which they would have to be producing to induce him to produce nothing — that is, provided that the output OM (or OR) in Diagram 105 is less than OS (or ON). Assumption (j) stated that each duopolist supposes that irrespective of the output plan which he may decide to implement in the current period, his rival will maintain his output at the same level as in the previous period. In other words, each duopolist behaves as if a change in his own output will not cause a change in the output of his rival: thus, in Diagram 105, A assumes that if he reduces his output from OM (its level in period 1) to Oa_2 (its level for period 2) B will continue to produce Ob_1. When an oligopolist acts on this kind of assumption about rivals' behaviour, we shall say (following Frisch)*

* Ragnar Frisch, *Monopoly — Polypoly — The Concept of Force in the Economy.* (Translated by W. Beckerman. International Economic Papers, Number 1.)

that he behaves *autonomously*. We have shown that an equilibrium will be reached if each oligopolist continues to act autonomously; we shall now examine the likelihood of their continuing to do so.

It is clear from Diagram 105 that the assumption which each duopolist makes about his rival's reaction is not being confirmed by events. Thus, in period 1, A expects B to produce zero, but B actually produces Ob_1; in period 2, B expects A to produce OM, but A actually produces Oa_2; in period 3, A expects B to produce Ob_1, but B actually produces Ob_3, and so on. We would therefore expect each duopolist to observe that changes in his own output are followed by changes in the opposite direction in the output of his rival. Whether or not he attributes a causal role to changes in his own output will depend on the length (in terms of calendar time) of each production period, and on the general stability of the market environment. If the production period is relatively long, he may not relate a current change in his rival's output to a change in his own output that occurred some considerable time ago; if demand and cost conditions vary appreciably from one period to another, he may ignore, because he cannot isolate or measure, the effect on his rival's output of changes in his own. If the production period is relatively short and the market environment relatively stable, each duopolist is likely to cease behaving autonomously and to seek some alternative and more 'correct' hypothesis about his rival's behaviour. Each duopolist might assume, for example, that changes in his rival's output are functionally related to changes in his own: thus, to take the simplest example, A might suppose that if he reduces his output from any level by 10 per cent, B will raise his output by 5 per cent. When A or B acts on this kind of assumption, he is said to act *conjecturally*.* If each duopolist acts conjecturally, then reaction curves can be drawn, and their point of intersection (or one of their points of intersection) may denote a position of stable equilibrium. This equilibrium, however, is no more likely to be attained in practice than that denoted by I in Diagram 105: if the reaction curves do not coincide with one another, then each duopolist will observe that his conjectures are not being fulfilled by his rival's actual behaviour, and each is therefore likely to seek some more 'correct' hypothesis about

* See R. Frisch, *op. cit.*

his rival's reactions. We shall explore further the nature and consequences of conjectural behaviour in the model which follows.

It is unlikely, therefore, that we shall ever find an actual oligopolistic market that is accurately described by the simple Cournot model wherein each oligopolist acts autonomously and output is the sole 'parameter of action'.* In this sense, the model is not 'useful'. It is useful, however, in the sense that it illustrates the distinguishing feature of an oligopolistic market, namely, the fact of mutual interdependence. During the progress (which we described on pages 326–7 above) towards the equilibrium denoted by *I* in Diagram 105, it is clear that a change in *B*'s output causes *A* to change his output, and that *A*'s reaction causes a further change in *B*'s output, and so on. We make this mutual interdependence clearer to us by making our duopolists unaware of it or so bemused by it that each of them ignores it when choosing his plans. The main justification for the assumption of autonomous behaviour is, then, the pedagogic usefulness of the model that is based on it. The Cournot model with output as the parameter of action suffices for this purpose. We shall not, therefore, examine the Bertrand model,† in which each oligopolist acts autonomously and price is the action parameter, nor shall we consider models of autonomous behaviour in which the oligopolists produce differentiated products and in which product quality or advertising expenditure is the parameter of action.

LEADERSHIP MODELS

We shall consider three leadership models, and a definition of leadership will emerge from the first of them. In our first model, we shall maintain the assumptions (*a*) to (*i*) inclusive, which we listed on page 321 above, and alter only assumption (*j*). In its stead, we shall suppose (1) that *A* conjectures that *B* will

* The term 'parameter of action' was first used by Frisch in the article mentioned in the footnote to page 327. An action parameter means any variable whose value lies within a firm's control. Thus, the sole parameter of action for a firm that operates under conditions of pure competition is the quantity of its output; the parameter of action for a monopolistic competitor is price or output, or product quality, or advertising expenditure.

† See Joseph Bertrand, Review of Cournot's *Recherches* . . . in *Journal des Savants*, September 1883. For a fuller consideration of the Cournot and Bertrand models, see William Fellner, *Competition Among the Few*, New York,

accept *A*'s output as a datum when he (*B*) is making an output plan, and (2) that *B* actually behaves in this way — that is, acts autonomously. Given assumptions (*a*) to (*i*) inclusive, we can, as before, draw the profit-indifference curves of each duopolist. Some of these are drawn in Diagram 106, Our assumption that *A* knows that *B* will act autonomously means that *A* knows *B*'s Cournot reaction function, which is shown by the line *RS*. This line shows us (and *A*) the output which *B* would plan to produce (and actually does produce) in each period at each level of *A*'s output. The points at which this line cuts *A*'s profit-indifference curves show the alternative profits per period that are open to *A* while *B* behaves in this way. Of these, *A* will choose that denoted by the point L_a where *B*'s reaction curve just touches one of his (*A*'s) profit-indifference curves: L_a will promise *A* the maximum profits per period, for any point either to the right or to the left of it on *RS* lies on a lower* iso-profit curve. *A* will therefore plan to produce an output of Oa_r per period, and *B* will produce Ob_r per period. In this model, *A* is the 'output-leader' and *B* the 'output-follower': *A* leads in that he chooses the output which he will produce in the light of his

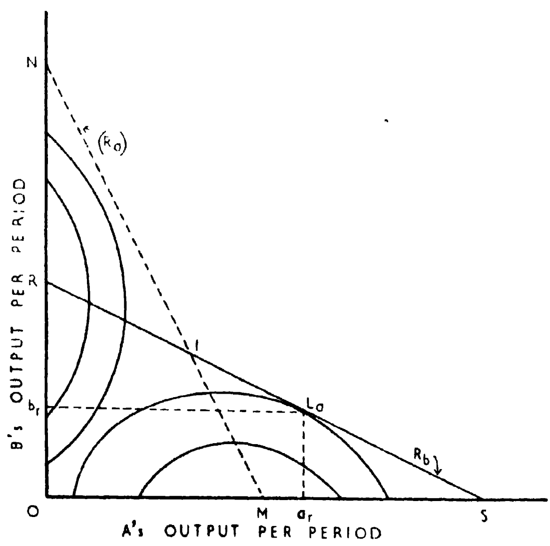


DIAGRAM 106

* Lower in terms of profit, but 'higher' in terms of position on the diagram.

(correct) conjectures about *B*'s reactions; *B* follows in that he accepts any output that *A* might produce as a datum. In the leadership model, the leader has no reaction curve, for he chooses that point on the follower's reaction curve which promises him the greatest profits. There is, therefore, no 'path' by which the leadership equilibrium will be reached, for the point L_a will be established immediately by *A*. When *A* acts (and is allowed by *B* to act) as the output-leader, his profits will be higher and *B*'s lower than they would have been had both *A* and *B* acted autonomously. In Diagram 106, the dotted line *MN* shows *A*'s reaction curve when he acts autonomously. The point *I*, at which the two reaction curves intersect one another, lies above L_a in *A*'s profit-indifference map and below L_a in *B*'s indifference map: that is, *A* will prefer L_a to *I*, and *B* will prefer *I* to L_a.

If the leadership equilibrium is to be maintained over a succession of periods, then *A* must be willing to accept *B*'s present pattern of reaction as shown by the curve *RS*, and *B* must remain ignorant of the fact that *A* knows his (*B*'s) reaction curve. It is clear from Diagram 106 that *A* would earn larger profits per period if he could force *B* on to a reaction curve that lay below *RS* and so touched a higher* iso-profit curve of *A*; and *A*, by threat and rumour, might try to persuade *B* to react along such a curve. If *B* suspects that *A* is aware that he (*B*) is acting autonomously, *B* may attempt to convince *A* that he will react along a curve that lies above *RS* and in such a position that the point L_a which *A* will choose on it lies on a lower indifference curve for *A* and on a higher profit-indifference curve for *B*. It is likely, therefore, that even in this simple leadership model, each duopolist will seek to alter to his own advantage the assumption which he thinks his rival is making about his reactions.

In our second model, we shall suppose that each duopolist is striving after leadership. We shall continue to make assumptions (*a*) to (*i*) which we listed on page 321 above. In addition, we shall suppose that each duopolist assumes that his rival will act autonomously. The probable consequences of these assumptions are illustrated in Diagram 107. We shall suppose that

* The adjectives 'higher' and 'lower' when applied in this and later paragraphs to an indifference curve refer to the value of the profits which it represents and not to its position in the diagram.

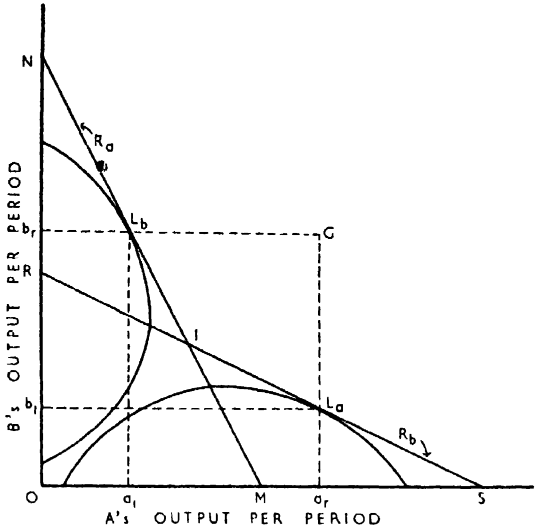


DIAGRAM 107

A has been a monopolist and that *B* suddenly and unexpectedly appears to compete with him. No *modus vivendi* has yet been reached, and each is laying his plans for period 1, the first period of their co-existence. On our assumptions, *A* believes that *B* will react along *RS*, so that he (*A*) will plan to produce Oa_r in period 1, expecting *B* to produce Ob_1; *B* assumes that *A* will react along *MN*, so he (*B*) plans to produce Ob_r, expecting *A* to produce Oa_1. In period 1, therefore, the total output of the commodity will be Oa_r plus Ob_r, or that denoted by the point *G*. Since *G* lies on a lower iso-profit curve in *A*'s map than L_a, and in *B*'s map than L_b, the profits which each duopolist earns in period 1 will be much less than what he expected to earn. In this way, each duopolist will discover that his rival is not behaving as he expected him to behave. During the periods which follow, each will seek some more 'correct' conjecture about his rival's reactions: he may do this by observing how his rival's output responds to experimental variations in his own; or he may try to force his rival to react along some reaction function that he prefers.

We see, then, that when each duopolist aspires to leadership, the hypothesis that each makes about his rival's behaviour will be proved wrong as soon as it is tested, and the assumptions by

which the model was defined do not help us to identify what new hypothesis each duopolist will choose.

This kind of model nevertheless helps us, for it can be used to illustrate the prelude to bargaining or collusion. Let us, for the sake of variety, take another member of the same family of models.* Let us suppose that (*a*) there are only two independent firms; (*b*) each produces and sells a product that is a close, but not a perfect, substitute for that of the other; (*c*) the product of each duopolist is perishable, so that in each period all that is produced must be sold; (*d*) there are many knowledgeable buyers of each firm's product; (*e*) the parameter of action is product quality, and this is measured by the weight (in milligrammes) of their respective products; (*f*) each duopolist knows the profits that he would earn for all values of his own and his rival's action parameter; (*g*) prices and advertising expenditures are data, and (*h*) each duopolist seeks the maximum profits per period. Given these assumptions, the profit-indifference map of each duopolist can be drawn. The iso-profit curves of *A* (and of *B*) will be as shown in Diagram 108, because (1) at any given

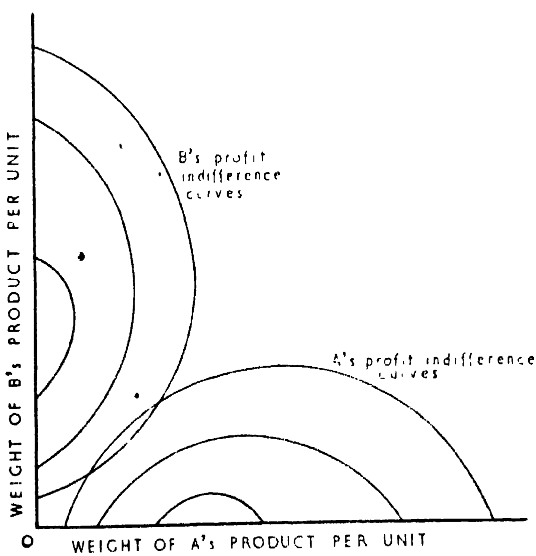


DIAGRAM 108

weight per unit of *A*'s product, *A*'s profits will decline as the weight of *B*'s product per unit is increased; (2) at any given weight per unit of *B*'s product, as the weight per unit of *A*'s product rises, *A*'s profits will rise for a while and then decline as the costs of producing the heavier product begin to outstrip the rise in the demand for it; and (3) the weight per unit of *A*'s product that promises him the maximum profits will be the greater, the greater is the weight per unit of *B*'s product.

Let us now suppose that *A* and *B* have agreed to meet to fix the weights of their respective products. For brevity's sake, we shall call the weight of *A*'s product per unit, x, and the weight per unit of *B*'s product, y. Neither duopolist will accept any combination of values for x and y that promises him profits which are less than those that he believes he could earn by acting independently of his rival. The maximum profits that *A* (or *B*) might expect to earn if no agreement is reached may be determined as follows. The profit-indifference map of *A* is drawn in Diagram 109(*a*). When there is no agreement, we shall suppose (taking the simplest case) that *A* believes that *B* will act autonomously — that is, that *A* assumes that *B* will react to changes in his (*A*'s) parameter along *RS*, which is *B*'s Cournot reaction curve when product-quality is the action parameter. *A* will therefore plan to fix the weight of his product at Ox_r, believing that *B* will choose Oy_r, and *A* will expect to earn the profits denoted by the iso-profit curve on which L_a lies if no agreement is reached. Similarly, we shall suppose that *B* acts conjecturally — that he believes that *A* will react along *MN* in Diagram 109(*b*): if no agreement is concluded, *B* will expect to earn the profits denoted by that one of his iso-profit curves on which lies L_b. The figures in Diagrams 109(*a*) and 109(*b*) are superimposed on one another in Diagram 110. Any combination of values for x and y that is likely to emerge from the negotiations must lie within the shaded area, which is enclosed by the iso-profit curves of *A* and *B* on which lie L_a and L_b respectively, for any point within this area lies on a higher iso-profit curve for each duopolist. *A* will not accept any combination of values for x and y such as that denoted by *F*, for since *F* lies on a lower one of his profit-indifference curves than L_a, it promises him profits which are less than those which he feels he can command by independent action. Similarly, *B* would not accept any combi-

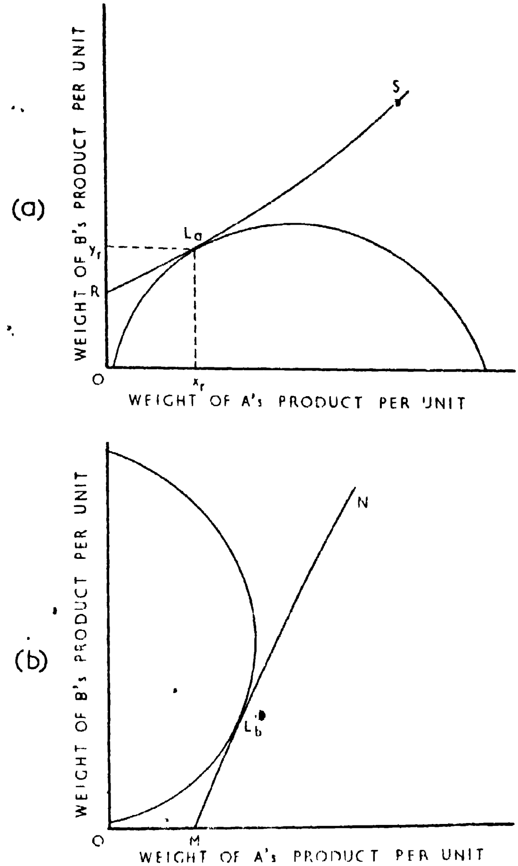


DIAGRAM 109

nation of values for x and y such as that denoted by G. Of the values for x and y that lie within the shaded area, some are more likely to be agreed upon than others: thus, if the negotiators begin by considering the values denoted by J, they are likely to discover that each duopolist could earn higher profits by accepting higher values for x and y, for by moving north-eastwards from J they will reach a higher iso-profit curve for each of them; similarly, if they are initially contemplating the values denoted by K, they are likely to discover that each duopolist could reach a higher iso-profit curve by moving south-westwards from K. It would seem, then, that if the two firms are roughly similar in size, resources and in the personalities of those who control them,

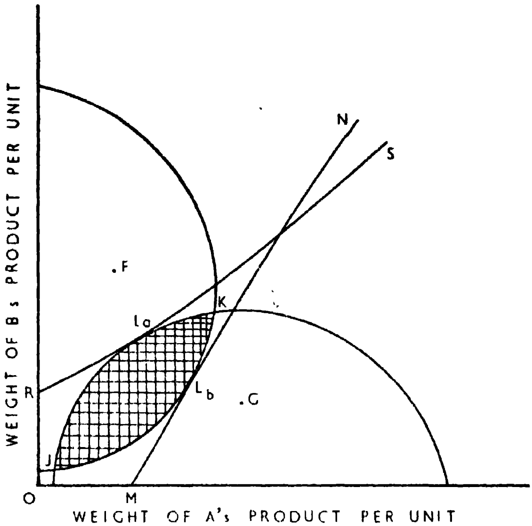


DIAGRAM 110

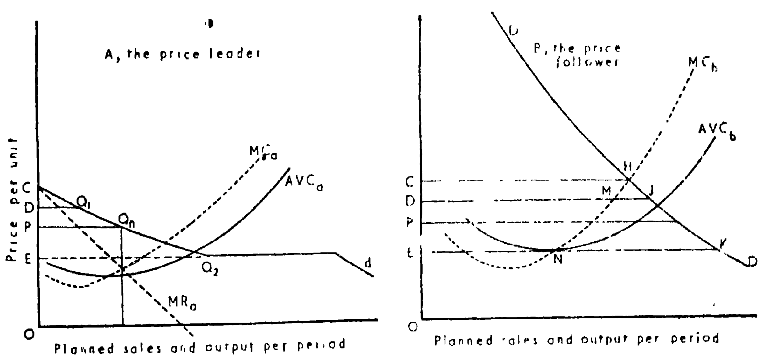
the values of x and y that they will agree upon will lie near the centre of the area of negotiation.*

A model in which each duopolist acts conjecturally may, therefore, help us to illustrate the limits within which bargaining may take place. If the duopolists accept the area of negotiation, then the hypothesis which each makes about his rival's behaviour may not be tested. The negotiations may break down if either duopolist questions the bargaining range (as shown by the shaded area in Diagram 110); if they do, then each must seek a new hypothesis about his rival's behaviour and reactions, for the old one will have been proved untenable by events. If either duopolist suspects the kind of conjecture that his rival is making about his reactions, he may attempt to alter it, for each will gain if he succeeds in lowering the bargaining limit of the other. It can be seen from Diagram 110 that when the iso-profit maps are given, the size of the area of negotiation (and therefore the extent of the increase in profits that agreement might promise to either participant) depends on the position of RS and MN. If A suspects that B expects him to react along MN, it will clearly be to A's advantage to persuade B that he (A) will react along a

* If the profit-indifference curves on which lie L_a and L_b do not overlap when superimposed on one another, then neither duopolist will be willing to negotiate, for each will believe that he can earn higher profits by independent action.

curve that lies to the right of MN; similarly, it will be to B's advantage to convince A that his (B's) reaction curve lies above RS. In that way, A (or B) might hope to convince B (or A) that he would earn lower profits if no agreement is reached.

Our third and last leadership model is one in which there is price-leadership. There is price-leadership when 'the price at which most of the firms in an industry offer to sell is determined by adopting the price announced by one of their number'.* In this model, we shall take the simplest example: we shall suppose that (a) there are two independent firms; (b) each produces a product that is a perfect substitute for that of the other; (c) the product is perishable so that in each period the output of each duopolist must all be sold; (d) there are many knowledgeable buyers of the product; (e) each duopolist knows the market demand curve for the product; (f) each seeks the maximum profits in each period; and (g) A assumes that B will always charge the same price as that which he (A) fixes, and B actually behaves in this way. Given these assumptions, we can illustrate the choice of a price by A, the price-leader. In Diagram 111(b), DD is the total demand curve for the product, and MC_b and AVC_b are the marginal and average variable cost curves respectively of B, the price-follower; in Diagram 111(a), MC_a and AVC_a are the marginal and average variable cost curves respectively of A, the price-leader. At each price which A fixes, B will offer for sale the quantity that promises him the greatest profits,



(a) (b)

DIAGRAM 111

and the amount by which the total quantity of the product that is demanded at that price exceeds *B*'s sales will be available for *A*. If *A* fixes the price at *OC* per unit, *B* will produce and sell *CH* per period — the output at which his marginal cost of production is equal to *OC*; at this price, *B* will be supplying all that buyers wish to buy, so that *A*'s sales will be zero. The point *C* in Diagram 111 (*a*) will therefore be one point on *A*'s 'conjectural demand or sales curve'. If *A* fixes the price at *OD*, *B* will offer *DM* per period, and *MJ* ($=DQ_1$ in Diagram 111 (*a*)) will be available for *A*; if *A* were to fix a price of *OE* per unit, *B* would offer *EN*, and *NK* ($=EQ_2$ in Diagram 111 (*a*)) would be left to *A*. At prices below *OE*, the whole of the market demand will be open to *A*. The curve CQ_2d in Diagram 111 (*a*), then, shows the quantity that the leader expects to be able to sell at each price. The leader will fix the price at the level which promises him the greatest profits — that is, at *OP*, where MR_a (the marginal revenue curve corresponding to CQ_2) cuts MC_a. By charging *OP* per unit, and satisfying the residual demand of PQ_n per period, the price-leader is maximising his profits within the limits set by the follower's behaviour.

We may develop variants of this model of price-leadership by increasing the number of oligopolists, by introducing differentiated products, and by positing other relationships between the leader's and the followers' cost curves than that shown in Diagram 111. These variants, however, add less to the economics of leadership than they do to its geometry. Price-leadership in some form* is common in actual oligopolistic markets.† It is probable that when it occurs it is based on some kind of agreement. It may be a result of tacit or implicit agreement: if there is one large firm, and several small firms, in an oligopolistic industry, the latter will be malleable to the wishes of the former, for they have little ability to inflict losses on the dominant firm and less capacity to bear the losses which it might inflict on them. The dominant firm need not consult with the smaller firms in order to establish its leadership: it need merely punish them if they do not accept it. If it forces its rivals to act as price-

followers, we must presume that that is the kind of reaction that is most profitable for the leader. If the oligopolists are more or less of the same size, then the agreement on which leadership is based is likely to rest on firmer foundations than tacit acceptance.

THE KINKED OLIGOPOLY DEMAND CURVE

In this model, we shall suppose that (*a*) there are several firms in an oligopolistic industry; (*b*) each produces a product that is a close substitute for that of each other firm; (*c*) product qualities are constant, advertising expenditures are zero, and some relationship between the prices of the differentiated products has already been established and is now obtaining; (*d*) each oligopolist believes that if he lowers the price of his product, his rivals will lower the prices of their products *pari passu*, and that if he raises his price, they will maintain their prices at their existing levels. Given these assumptions, each oligopolist will believe that the relationship between his price and his sales will be similar to that shown by the curve *dPD* in Diagram 112:* if he were to raise his price above *MP*, he would expect his sales to fall off markedly, for his product would become relatively dearer; if he were to lower his price below *MP*, he would

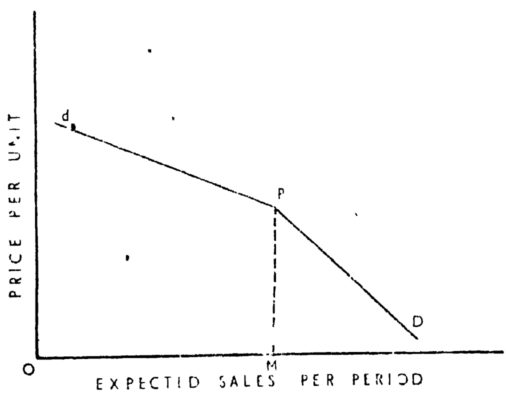


DIAGRAM 112

expect no appreciable increase in his sales, for his product would be prevented from becoming relatively cheaper by the price reductions of his rivals. We shall call *dPD* the oligopolist's 'conjectural demand or sales curve', for it shows the relationship between his price and his sales given his conjecture about the reactions of his rivals. The position of this curve is defined by the location of *P*, and the co-ordinates of *P* are *MP*, the price at which the oligopolist now happens to be selling his product, and *OM*, the quantity of it that he is currently selling in each period. The price *MP* is a datum, and not a quaesitum, for this model. An oligopolist is more likely to make assumption (*d*) above, if the price *MP* has been fixed by some informal agreement or by a rival who is accepted as the price-leader; in these circumstances, the assumption will reflect each oligopolist's assessment of the penalties that his rivals will inflict on him if he tries to act independently. If the price *MP* is a result of an explicit agreement, we are unlikely to find a kinked demand curve for the individual oligopolist, for there will probably be specified and known penalties for deviating from it.

The price *MP*, however fixed, is not necessarily inconsistent with the maximisation of profits. In Diagram 113, *MC* is the oligopolist's marginal cost curve, and *MR* is the marginal revenue curve corresponding to *dPD*; since the former cuts the latter vertically below *P*, the price *MP* is that which promises the greatest profits per period. With given marginal costs of pro-

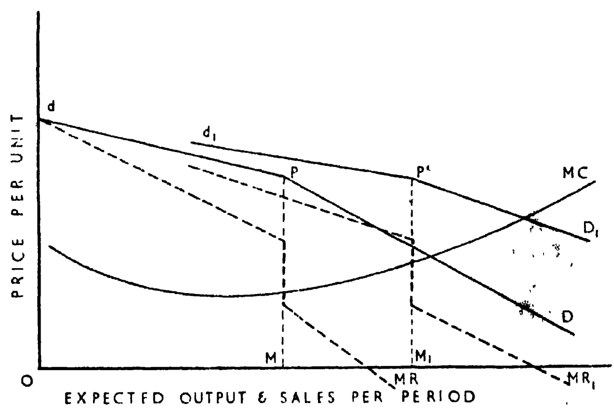


DIAGRAM 113

duction, MP is the more likely to be the profit-maximising price, the longer is the vertical portion of the marginal revenue curve. The length of the 'discontinuity' depends on the relative elasticities of demand at P of the curves dP and PD. This can be easily confirmed: we know (see *supra*, page 285) that marginal revenue at MP is equal to $MP(1 - 1/e)$; it follows that the greater is the elasticity at P on dP, and the less is the elasticity at P on PD, the greater will be the differences between the two marginal revenues, and the greater therefore will be the length of the discontinuity. The elasticity of dP will reflect the degree to which the oligopolist's product can be substituted for that of his rivals; in the extreme case, if the oligopolists' products are homogeneous it will be perfectly elastic. The elasticity of PD will reflect the elasticity at each price of the 'demand' for the class of product that the oligopolists are producing.

While the hypothesis about rivals' reactions that gives us the kinked demand curve does not explain why the price is at its present level, it does explain why the price might remain stable at that level as time passes. If the demand for the product of oligopolist A (who makes this hypothesis) rises, he will become aware of it by an increase in his sales. Since the assumption that he makes about the probable reactions of his rivals is in no way dependent on the level of his sales, an increase in demand will not, *per se*, induce him to seek an alternative hypothesis: that is, A will interpret a rise in the demand for his product as a rightward shift of the dPD-curve to d_1PD_1 in Diagram 113. This maintenance of the price at MP in the face of a rise in demand is not necessarily inconsistent with the maximisation of profits. In Diagram 113, we have assumed that the dPD and d_1PD_1-curves are iso-elastic, and the marginal cost curve MC cuts the marginal revenue curve (MR_1) corresponding to the new demand curve within the discontinuity, so that M_1P ($= MP$) continues to be the price at which the oligopolist's profits will be maximised. Similarly, it can be shown that the profits of oligopolist A may still be greatest at the price MP, even after his costs of production have risen. Thus, in Diagram 114, the marginal cost curve rises from MC to MC_1; since MC_1 cuts the marginal revenue curve corresponding to dPD within the discontinuity, his profits in the new cost conditions will be maximised by keeping his price at MP. If the increases in demand and costs

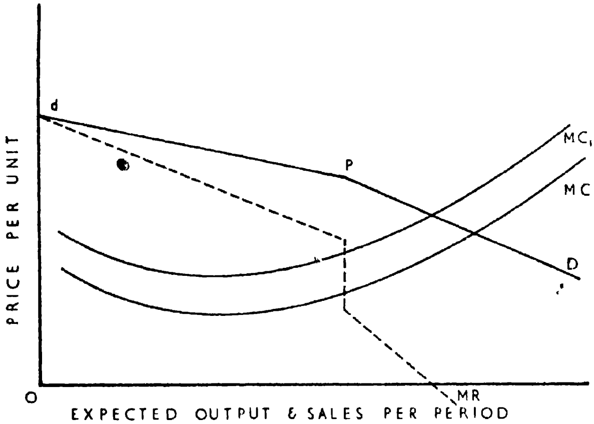


DIAGRAM 114

are not confined to a single oligopolist but affect all firms producing that class of product, then it is likely that a new informal agreement will be made or that the price-leader will adjust his price; if all the oligopolists are affected by the increases in demand and costs, we would expect the price of their products to rise.

COLLUSIVE OLIGOPOLY*

In the simple Cournot model, and in the models in which each firm aspired to leadership, there was no agreement whatsoever between the oligopolists. In the models in which one firm was accepted as the price- or output-leader, there was agreement between the oligopolists on the method by which the value(s) of the action parameter(s) should be fixed. In the kinked demand curve model, we supposed that the firms had already agreed upon a set of values for the prices of their products, and we explored one way in which these values might be maintained. In this, the last section on oligopoly, we shall discuss the economics of agreements between the firms in an oligopolistic industry. We shall first suppose that the oligopolists have agreed to extract the *maximum maximorum* of profits per period from the market(s) for their product(s), and we shall illustrate the choice of values for the variables under their control which promises to

achieve this aim. We shall then examine how and why they may be prevented from achieving this objective or deterred from pursuing it.

Initially, we shall suppose that (*a*) there are only two firms in the oligopolistic industry; (*b*) each produces and sells a product that is a perfect substitute for that of the other; (*c*) the product is perishable; (*d*) there are many knowledgeable buyers of the product; (*e*) each knows the market demand for the product; (*f*) the two firms have different cost curves; (*g*) each firm has the same expectations about the prices and productivities of the productive services which they use; (*h*) the price of the product is the sole parameter of action of each firm; and (*i*) they are contemplating whether or not to agree upon a value for the price that will promise the *maximum maximorum* of profits per period to both of them jointly. The choice of this price, and its implications, are illustrated in Diagram 115. The average and marginal costs of production of duopolist A are shown by AC_a and MC_a respectively in figure (1); AC_b and MC_b in figure (2) show the average and marginal costs respectively of B; the DD-curve in figure (3) is the market demand curve for the product. When A and B are jointly earning the *maximum maximorum* of profits per period from the production of the product in their respective plants and its sale in their common market, they will together be producing the output at which marginal cost and marginal revenue are equal to one another (for that is merely another way of saying that profits are being maximised), and this total out-

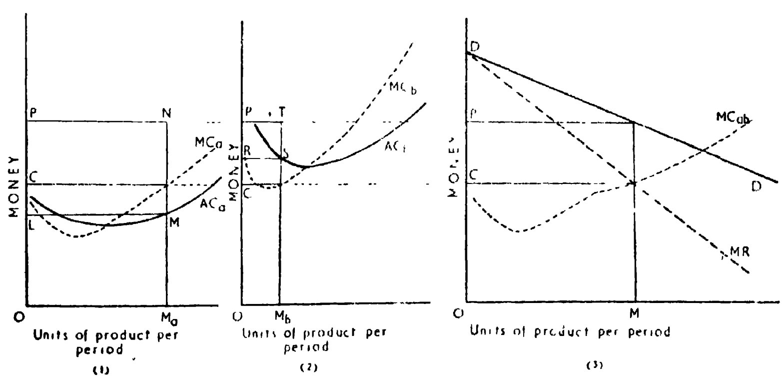


DIAGRAM 115

put will be distributed between their respective factories in such a way that the last unit produced by each adds the same sum to the costs of production of each, for if that is not so then the sum of their total costs can be lowered by transferring output from A to B, or vice versa.

In Diagram 115, we can discover geometrically the price of the product and the outputs of A and B at which these conditions will be fulfilled. The MC_{ab}-curve in figure (3) is obtained by adding together laterally the MC_a- and MC_b-curves: thus, if OC is the marginal cost of the OM_a-th unit in A's plant and of the OM_b-th unit in B's, the co-ordinates of the corresponding point on the MC_{ab}-curve will be OC and OM_a plus OM_b ($= OM$). This MC_{ab}-curve shows the minimum addition to total costs that will be incurred by producing the last unit of any output, and this illustrates the second condition mentioned above. The first condition is fulfilled when the total output of the product is OM per period and its price OP per unit, for at that output and price marginal revenue and the (minimum) addition to costs of production from producing the OM-th unit are the same. The duopolists will therefore plan to sell OM units of their product at a price of OP per unit, and A will produce OM_a and B, OM_b per period.* The output OM and the price OP are the 'monopoly' output and price respectively: that is, they are those that would be produced and charged respectively were A and B to merge together to form a single firm that operated two plants in which the costs of production were as shown in figures (1) and (2) in Diagram 115. For brevity's sake, we shall hereafter say that the price OP and the output OM (distributed between A and B in the proportion $OM_a : OM_b$) define the 'monopoly' solution.

The maximisation of their joint profits requires that A and B produce OM_a and OM_b per period respectively. This distribution of the 'monopoly' output OM implies a distribution of profits: A may expect to earn profits of $LMNP$ per period and B of $RSTP$ per period. The sum of $LMNP$ and $RSTP$ will be greater than the sum of the profits that A and B would earn with any distribution of output between them at any price other than OP, or with any other distribution of the output OM between them at the price OP. While the 'monopoly' solution promises the

* Since the MC_{ab}-curve was obtained by adding together the outputs of A and B at each level of marginal cost, OM_a plus OM_b must together be equal to OM.

maximum joint profits to A and B, however, either might feel that he could command a larger profit* by acting independently of his rival. Thus, A might believe that, in the absence of any agreement, B's reactions to changes in his (A's) parameter(s) will promise him profits per period greater than $LMNP$. In these circumstances, A will not accept the distribution of output which the monopoly solution dictates unless some device is found for divorcing the profits which he receives during the period from the profits which he earns when producing his share of the monopoly output. Many such devices are possible: for example, A and B might pay the profits which they earn, when producing outputs of OM_a and OM_b respectively and selling them at a price of OP per unit, into a central pool or fund from which each then receives a sum which is not less than that which he believes (and which his rival agrees) he could earn by acting independently. A pooling agreement of this kind will always make it possible for A and B to maximise their joint profits, provided that the sum of the profits which they believe they can earn by independent action does not exceed the monopoly profits.† If the sum of the profits which each believes he can earn by acting independently of his rival exceeds the monopoly profits, then no agreement is possible, for the expectations of A and B are inconsistent with one another.‡ In the ensuing periods, each duopolist will test his hypothesis about his rival's behaviour, and as its incorrectness becomes manifest will be forced to revise it. When the sum of the profits that each expects to earn if there is no agreement again falls short of the monopoly profits, then the 'monopoly' solution will again appear attractive.

Thus far we have confined our analysis to the model defined on page 343 above. We have so far assumed that the duopolists have different costs of production in producing the same product and that they have identical expectations about the behaviour of the demand for their product and the supplies of the productive services they employ during the period that lies

* That is, profits which are larger than the share of the 'monopoly' profits which he would earn when producing his share of the 'monopoly' output.

† This proposition can be interpreted in terms of Diagram 110. If the sum of the profits which A expects to earn at L_a and which B expects to earn at L_b falls short of the monopoly profits, there will be an overlap area between the iso-profit curves on which lie L_a and L_b respectively.

‡ In this case, in terms of Diagram 110, the iso-profit curves on which L_a and L_b lie will not overlap with one another.

ahead. The emergence of a unique and agreed 'monopoly' solution depends on this latter assumption. If A believes that the demand for the product will rise and that the prices of the productive services will fall, and if B expects demand to fall and costs to rise, then A will expect the monopoly profits (and his share of them) to be relatively large and B will expect them to be relatively small. In these circumstances, ignoring the possibility that either A or B might deem unacceptable his (earned) share of his estimate of the monopoly profits, there is little likelihood of agreement between them. Agreement may become the more likely as the gap between their expectations narrows as events confirm the estimates of A (or B), and thus lead B (or A) to revise his estimates of the future behaviour of demand and costs.

Our analysis and the conclusions it yields require little modification if the model is extended to reflect more accurately actual oligopolistic markets. Let us now suppose (*inter alia*) that (*a*) there are several firms in the oligopolistic industry; (*b*) each produces and sells a product that is a close substitute for that of each other; (*c*) each has the same expectations about the behaviour of the 'demand' for the product-group that they are producing and of the supplies of the productive services that they are using, and (*d*) they are considering the implications of an agreement that would promise the *maximum maximorum* of profits per period to all of them jointly. In this model, each firm will have many parameters of action: the present and future profits of firm A, for example, will depend on the relative values that he attaches to the price of his product, its quality, his techniques of production, and his advertising expenditure, and to his expenditure on the search for new variants of the product, new methods of production, and new kinds and avenues of advertisement. There will still exist a 'monopoly' solution, however, in this more complex model, for there will be some set of values for the parameters of action of the oligopolists which promises them jointly the *maximum maximorum* of profits. As before, each firm might consider unsatisfactory the profits that it might expect to earn when its parameters had the values dictated by the 'monopoly' solution, and the chances of this happening become the greater the larger is the number of firms in the oligopolistic industry and the larger is the number of variables whose values lie within the control of each of them. In this,

as in the simple model, a pooling agreement will still make possible the maximisation of the joint profits, provided that the sum of the minimum profits that each firm would demand from the pool is not greater than the 'monopoly' profits. When the parameters include research and development, and variations in product-quality and advertisement, however, it is probable that the sum of the expected profits from independent action will exceed the monopoly profits: for competition *via* variables other than price requires more skill* (and perhaps more 'luck') than competition through price, and each firm may tend to overestimate its proficiency in non-price competition and the good fortune that it expects to attend its efforts in that direction. If the oligopolists have different expectations about the future behaviour of demand and costs, each will have his own notion of what constitutes the monopoly solution, and some alternative though less profitable agreement must be sought.

We may conclude, then, that while the 'monopoly' solution promises the *maximum maximorum* of profits to the firms in an oligopolistic industry, it may not be reached for any one of three reasons: (*a*) because the oligopolists have different expectations about the future behaviour of demand and costs, and therefore about what constitutes the 'monopoly' solution; (*b*) because, while agreeing upon the monopoly solution, the sum of the profits that they expect from independent action exceeds the expected 'monopoly' profits; and (*c*) because, in the absence of reasons (*a*) and (*b*) above, the oligopolists will not accept (or are prevented from accepting) the pooling agreement which makes it possible for each to receive a share of the 'monopoly' profits that he deems satisfactory. If the oligopolists are deterred by any one of these reasons from effecting an agreement to maximise their joint profits, they need not necessarily eschew collusion of any kind. They may seek other agreements which are less comprehensive in that they do not cover all the variables whose values determine the distribution of profits between the firms, and which are potentially less profitable to the oligopolists taken as a group. We shall now examine briefly a few of these alternatives.

First, the oligopolists may agree to share the market. Let us

* More skill to inaugurate, though perhaps no less skill to carry to a successful conclusion.

again make the assumptions (*a*) to (*g*) that are listed on page 343 above, and let us consider the implications of an agreement between *A* and *B* to share the market for their product in the proportions 2 : 1. The market-shares that are agreed upon will be those that promise each duopolist a sum of profits per period that is not less than that which he believes he could earn either without any agreement or with any other kind of agreement. In Diagram 116, the average and marginal costs of production of *A* are shown by AC_a and MC_a respectively in figure (1); AC_b and MC_b in figure (2) show the average and marginal costs respectively of *B*, and the *DD*-curve in figure (3) is the market demand curve for the product. The D_aD_a and D_bD_b curves in figures (1) and (2) respectively are the market-share curves of *A* and *B*, and in our example, the permissible sales of *A* at any price on D_aD_a will be twice those of *B* at the same price on D_bD_b. Given the market-shares, *A* will plan to sell OM_a per period at OP_a per unit, for these plans promise him the maximum net revenue, and *B* will plan to sell OM_b at OP_b per unit. These plans cannot be simultaneously fulfilled, for since *A* and *B* produce the same product they must charge the same price. Having agreed to share the market, *A* and *B* must therefore agree upon a price for their product in the range OP_a to OP_b. The choice of a price and of actual values for the market-shares must be made simultaneously, for the profits that each duopolist can hope to earn if the agreement is effected will depend on both of these things.

If the duopolists agree upon a price of OP_n per unit, it is clear from Diagram 116 that each will be tempted to produce and

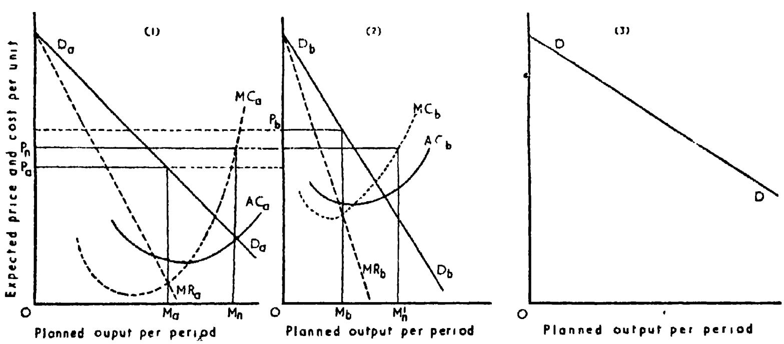


DIAGRAM 116

sell more than his share of the market, for by doing so he will increase his profits or diminish his losses. Thus, at OP_n per unit, *A*'s profits will be greatest when he is producing and selling OM_n per period, and *B*'s when he (*B*) is selling OM'_n per period. If both firms succumb to this temptation, then both will accumulate stocks of the product, and these in turn may tempt one or other to dishonour the agreement by reducing his selling price. If one firm succumbs, and successfully sells more than his allotted share at the price OP_n, then the other firm's sales (and its share of the market) will be *pro tanto* reduced. In recognition of these temptations, the simple market-sharing agreement is normally fortified by a system of fines and compensations: firms that exceed their allotted quotas must pay a proportional or progressive tax on their excess sales, and the proceeds are used to compensate the firms that are thereby prevented from fulfilling their quotas.

A market-sharing agreement is perhaps most likely to occur when the oligopolists incur different costs of production in making the same product, and when a pooling agreement (without which the 'monopoly' solution would be unacceptable) is illegal. Once made, the agreement will persist for as long as the oligopolists are satisfied with the market- (and profit-) shares that it promises, and these shares depend on the profits that the oligopolists believe they can earn by the most attractive alternative agreement or with no agreement. Even when the oligopolists are producing the same product, over the long period each may spend money on research into new methods of production or new variants of the product, and as these efforts are attended by different degrees of success, the acceptable market- and profit-shares will alter. When this happens, the existing agreement will be terminated, and replaced by one in which the market- and profit-shares are different, or by an agreement of a different kind.

This simple model by which we have illustrated the market-sharing agreement may be extended to include other parameters of action. When the oligopolists are producing products that are close substitutes for one another, the profits that each can command over any span of future periods will depend on the relative values of his price, product-quality, techniques of production, advertisement, and expenditure on research. In these

circumstances, there will exist some set of values for these variables that will distribute the 'market-demand' for the class of product that the firms are producing (and therefore profits) in any given proportions between them. This more inclusive market-sharing agreement will be subject to the same strains and stresses and will require the same safeguards in the way of penalties and compensations as the simple agreement that we have already examined. It is unlikely, however, that any such inclusive agreement will be reached, and for two reasons. First, the choice of a set of values for the relevant variables that is acceptable to all the participating firms may be impossible, for it rests not so much on ascertainable and measurable facts as on judgements about the future consequences of present changes in the relationship between prices, product-qualities, advertisements or research expenditures. Second, even if this choice is made, it may be impossible to devise a system of fines and compensations to safeguard the agreement, for changes in variables other than price are more easily concealed and their consequences are often less clear. For these reasons, when the number of variables is large, the oligopolists may agree on values for only one of them.

The variable that is most commonly the subject of agreement is price, for a reduction in price by one firm will usually have more immediate and marked effects on the sales of its rivals than an increase, for example, in its advertising or research expenditure. The price-agreement may specify the exact or minimum price that each oligopolist must charge for his product, or it may define the method by which the prices of the competing products must be fixed. The agreement may set out a uniform procedure that each firm must follow when fixing its price: thus there may be a table of 'standard', 'normal' or 'typical' costs and each oligopolist is obliged to base his price on these rather than on his own costs. Alternatively, if the oligopolistic industry consists of one large firm and several small firms whose costs are not very dissimilar, the choice of a price may tacitly be left to the former: the small firms might feel that it (the large firm) is the more likely to have a clear notion of the demand for the product or product-group and of costs of production than they have, and that it is therefore more likely to fix a price that approximates to the 'monopoly' level. In the price-leadership

model (*supra*, pages 337–9), we have described the choice of a price by the leader. In the kinked demand curve model (*supra*, pages 339–42) we have described one way in which a price-agreement might be maintained, without any explicit penalties or policing. A price-agreement is subject to the same stresses as any of the agreements that we have already examined: the agreement will generally disintegrate when one or more of the participants are convinced that he or they could command higher profits without it than within it.

In all the models of collusive oligopoly that we have examined so far, we have supposed that whether or not a particular agreement is reached depends simply on whether or not it promises each oligopolist a higher rate of profit than that which he could earn without it. This assumption though crude was useful while our purpose was simply to catalogue some of the different kinds of agreement that might occur and to adumbrate the circumstances in which each was likely to appear. However, if we wish to explain how the spoils that any agreement promises are shared between the participants — that is, what determines the distribution of the 'monopoly' profits or the relative market-shares — then this assumption must be refined.

Let us return to the model that is defined on page 343 above. If there is no agreement, the profits per period that duopolist *A* might expect to earn will depend on the hypothesis which he makes about the expected reactions of his rival. For each hypothesis that he might make, there will be an expected rate of profit.* If *A*'s objective is to earn the maximum profits per period, then he will only accept the agreement if his share of the 'monopoly' profits is not less than the maximum rate of profit that he believes he can earn without it. The share of the 'monopoly' profits that *A* will obtain, however, depends not only on his (*A*'s) estimate of his prowess if no agreement is reached; it depends also on his rival *B*, for *B* will only enter the agreement if it promises him a higher rate of profit. If neither *A* nor *B* questions his rival's estimate of the rewards of independent action, and if the sum of these rewards is less than the expected 'monopoly' profits, then agreement is possible, and the

* Strictly, since *A* does not know how his rival will react there will be a range of probable values for his profits for each hypothesis about his rival's behaviour. For simplicity's sake, we shall assume that he reduces this range to a 'certainty-equivalent' or that he acts as if he does.

precise terms of the agreement will depend on how the amount by which the 'monopoly' profits exceeds the sum of the minimum demands of A and B is divided between them. We shall define the 'relative strength' of an oligopolist as his power to command profits within an agreement, and we shall suppose that it is measured by the proportion of the joint profits which he obtains. We may then say that the outcome in our present example will reflect the relative strengths of the firms that participate in the agreement.

We have so far supposed that both duopolists accept the 'bargaining range' as defined by the maximum profits that each believes he could earn without the agreement. This is not likely to be generally true, for the agreement could be made potentially more profitable for either duopolist if he successfully lowered the bargaining limit of his rival. Thus, if the 'monopoly' profits are 100, the minimum demands of A and B 40 and 20 respectively, and their relative strengths in the proportion 3 : 1, then when each accepts the bargaining range, A will obtain 70 per period and B 30 per period; if B can lower A's estimate of the maximum profits that he (A) could earn without the agreement from 40 to 20, then, *ceteris paribus*, the agreement will promise A only 65 and B 35 per period; if A lowers B's minimum requirement to 10, then A will obtain $77\frac{1}{2}$ and B only $22\frac{1}{2}$ per period. Either duopolist can attempt to lower the bargaining limit of his rival by inducing him to revise the hypothesis on which his existing estimate is based: thus, in terms of Diagram 110, A can lower B's bargaining limit by shifting the point L_b eastwards, southwards, or with any degree of south-eastwardness in the diagram, and he may seek to do so by propaganda and rumours whose purport is that for any given value of B's parameter he (A) will give a much higher value to his parameter than B now expects; similarly, B may make the agreement potentially more profitable to him by convincing A that L_a lies north, west, or north-west of the position in which A now believes it to be.

It is clear, then, that if agreement is possible, its terms will reflect the relative strengths of the oligopolists who are parties to it. The relative strength of a firm, as we have defined it, will depend on the size of the profits which it believes it could earn if no agreement is reached, and on its power to depress the bargaining limits of its rivals. The estimated profits from indepen-

dent action will depend on certain objective and measurable characteristics of the firm and on the personality of the entrepreneur who guides it. Amongst the former, we must list the brute size of the firm, the nature of its liabilities and assets structures, and the shape and position of its cost function. If the firm is relatively large,* if a relatively large proportion of its assets is in the form of money or near-money, if the ratio of contractual liabilities (for example, debentures) to total liabilities is relatively low, and if its average total costs of production are relatively low and rise relatively slowly as its output is expanded, then, *ceteris paribus*, we would expect it to be able to command relatively large profits if no agreement is reached. Given all these facts, however, the actual profit-estimate on which the entrepreneur decides whether or not to enter an agreement will reflect his skill *qua* entrepreneur and his attitudes towards his rivals and the uncertainty that the future holds. These attitudes are in part inherited from his ancestors, and in part they are the consequence of the character of the development of his firm and of the history of its industry. There is little that can be said about an entrepreneur's ability to depress the bargaining limit of his rival(s), other than that it will reflect his skill as an entrepreneur and as a negotiator. Of all the determinants of relative strength that we have listed, it is probable that the objective factors and the entrepreneur's skill *qua* entrepreneur are the most important, for it is these that will shape the outcome if no agreement is reached. Negotiating skill and psychological attitudes can achieve more favourable results than the objective factors warrant only for so long as all the firms are unwilling to submit the hypotheses on which their bargaining limits are based to empirical testing.†

BILATERAL MONOPOLY‡

Bilateral monopoly exists when one buyer faces one seller. We shall take barter between two parties as our prototype of this

* Where relative size is measured by the proportion of the total output of the oligopolistic industry that the firm would produce at any given price for the industry's product or any typical set of prices for the industry's products.

† We have not so far considered oligopsony, nor shall we do so. For each of the models of an oligopolistic industry that we have examined in this chapter, we can construct a similar model of an oligopsonistic industry simply by substituting 'productive service' for 'product' and 'buyer' for 'seller' in each of the assumptions that define it.

‡ Our general approach in this section is similar to that of E. Schneider in

market structure. Let us suppose that (*a*) '*A* has . . . a basket of apples, *B* a basket of nuts', and that '*A* wants some nuts, *B* wants some apples';* (*b*) all *A*'s apples, and all *B*'s nuts, are homogeneous; (*c*) we are given the indifference maps of *A* and *B*; and (*d*) each party seeks to maximise his satisfaction. The consequences of these assumptions are illustrated in Diagram 117. *A*'s indifference map is drawn in figure (1) and this shows *A*'s tastes for apples and nuts and his preferences as between different combinations of them. *A*'s basket contains *OR* apples, and the indifference curve A_0 on which *R* lies divides all combinations of nuts and apples which *A* would prefer to *OR* apples from those he would deem less attractive. Similarly, the indifference curve B_0 in figure (2) illustrates *B*'s bargaining limit, for he will not trade with *A* unless it leads to a combination of nuts and apples which he prefers to any combination lying on B_0. In figure (3), *B*'s indifference map, after having been rotated anti-clockwise through 180°, is superimposed on *A*'s: $OZ (= OR)$ shows the number of apples in *A*'s basket and $O'Z (= OS)$ the number of nuts in *B*'s basket. *A* and *B* will only be willing to trade with one another if as a result of trade each is left with a combination of nuts and apples that lies within the area bounded by A_0 and B_0 in figure (3).

The assumptions that we have made so far are not sufficient to enable us to decide what quantities of nuts and apples will be exchanged.† They merely tell us that the point denoting these

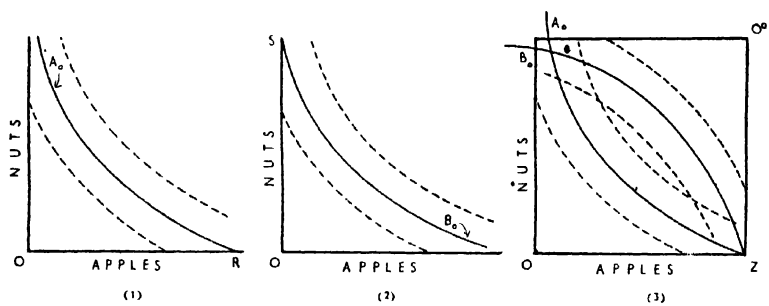


DIAGRAM 117

quantities must lie within the area enclosed by A_0 and B_0 in figure (3). If we wish to narrow the range of possible outcomes in this example of barter exchange, we must make some assumption about the market behaviour of A and B. We shall suppose initially that A and B do not enter into explicit negotiations with each other in pursuit of a mutually acceptable solution. In this context, we shall explore the consequences of the following three assumptions: first, that both A and B are price-takers; second, that A (or B) is a price-maker and B (or A) a price-taker; and third, that both A and B try to be price-makers. When that has been done, we shall illustrate the process of negotiation and describe its probable consequences.

The assumption that A and B are price-takers may be stated in other words: we may say that each is a quantity-adjuster, or that each behaves as if he were a pure competitor, or that each acts autonomously. The consequences of this assumption are illustrated in Diagram 118. The indifference map of A is drawn in figure (1). The slope of the straight lines radiating from R illustrate alternative prices for apples in terms of nuts: if A could buy OL/OR nuts for each apple, he would maximise his satisfaction by selling CR apples for OD nuts and thus acquiring the combination of apples and nuts denoted by the point P at which the price-line RL is a tangent to one of his indifference curves. When all points such as P are joined together, we have the curve RR', which is A's price-consumption or offer curve. The RR'-

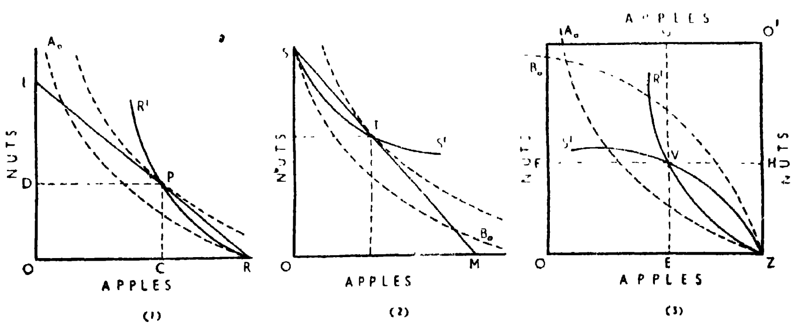


DIAGRAM 118

number of nuts that will be exchanged for one apple, and this will be equal to the quantity of nuts that B sells (A buys) divided by the number of apples that B buys (A sells). Similarly, the price of nuts in terms of apples will be equal to the quantity of apples divided by the quantity of nuts.

curve shows us the quantity of apples that *A* would be willing to sell at each rate of exchange between nuts and apples. The *SS'*-curve in figure (2) is derived in a similar way, and has a similar meaning. In figure (3), figures (1) and (2) have been superimposed on one another. The quantities of nuts and apples that will be exchanged and the rate at which they will be exchanged are implicit in the point *V* where the offer curves intersect: the price of apples in terms of nuts is shown by the slope of *ZV*, and at this price *A* will sell *ZE* apples for *OF* nuts, and *B* will sell *ZH* nuts for *O'G* apples. This represents an equilibrium position, for *OF* = *ZH* and *ZE* = *O'G* — that is, the planned purchases and sales of each commodity are the same. This model of bilateral monopoly is analogous to the simple Cournot model of oligopoly which we described at the beginning of this chapter: the offer curves are the analogues of the reaction curves, and our model enjoys the same advantages and suffers from the same defects as the Cournot model.

The equilibrium denoted by *V* in figure (3) of Diagram 118 can be illustrated in terms of demand and supply analysis. In Diagram 119, we measure the rate of exchange between nuts and apples on the vertical axis, and the quantity of apples demanded and supplied on the horizontal axis. Implicit in the *RR'*-curve in Diagram 118, there is a relationship between the rate of ex-

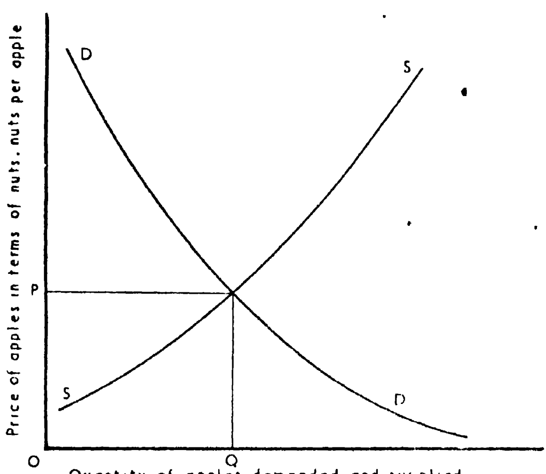


DIAGRAM 119

change and the number of apples that A would be willing to supply, and this relationship is shown explicitly by the SS-curve in Diagram 119. Similarly, DD is the demand curve for apples and it is derived from the SS'-curve in Diagram 118. The price of apples in terms of nuts will tend towards QP (= the slope of ZV in Diagram 118), and the quantity of apples that will be demanded and supplied at this price will be OQ (= ZE or $O'G$ in Diagram 118).

We shall next suppose that A is a price-maker and B a price-taker: that is, that A behaves as if he were a monopolist for apples and a monopsonist for nuts and B as if he were a pure competitor, or that A acts conjecturally and B autonomously. The consequences of this assumption are illustrated in Diagram 120. The ZA_0 and ZB_0 curves have the same meaning as in the preceding diagrams, and ZS' is B's offer curve. If A knows the quantity of nuts that B will sell (or the quantity of apples that B will demand) at each rate of exchange which he (A) might fix,* he will choose that exchange rate which promises him the maximum satisfaction. This is shown in the diagram by the slope of the straight line ZL_a, where L_a is the point at which B's offer curve is tangential to one of A's indifference curves. The equilibrium price of apples in terms of nuts will be UL_a/UZ, and at this price A, the monopolist, will plan to sell UZ apples; alternatively, we may say that the equilibrium price of nuts in terms of apples will be UZ/UL_a, and that at this price A, the monopsonist, will plan to buy OJ nuts. It is apparent that this model is analogous to that in which A acts (and is allowed by B to act) as the output-leader (see *supra*, pages 329–31), and it has similar merits and defects.

The information portrayed in Diagram 120 can be represented in terms of demand and supply analysis. In Diagram 121(a), DD is B's demand curve for A's apples and it is derived from B's offer curve; SS is A's 'marginal cost' curve for apples and it is obtained from the ZR'-curve in Diagram 118(3). Since A acts as a monopolist, his equilibrium will be implicit in the point E at which the marginal revenue curve corresponding to DD cuts SS: that is, A will plan to sell OC apples at a price of OD nuts per apple. The price OD is equal to UL_a/UZ and OC is the same as UZ in Diagram 120. In Diagram 121(b), the equili-

* That is, if A knows B's offer curve.

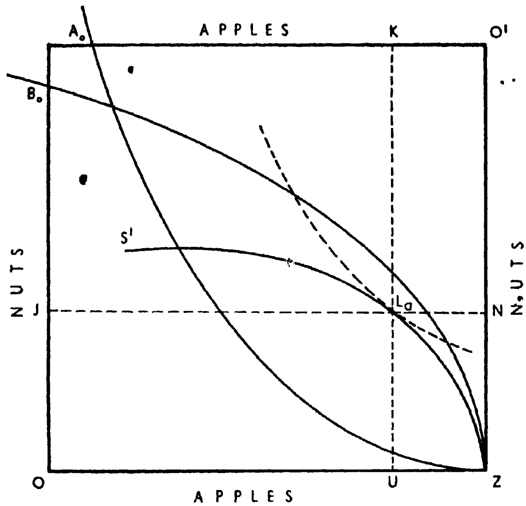


Diagram 120

brium of A *qua* monopsonist is shown: on the vertical axis, we measure the price of nuts in terms of apples, and on the horizontal axis, the planned purchases by A of nuts. In figure (*b*), the $S'S'$-curve gives the same information as the DD-curve in (*a*), and the $D'D'$-curve in (*b*) corresponds to the SS-curve in (*a*). If A acts as a monopsonist, his equilibrium will be implicit in the point H, where the marginal curve corresponding to $S'S'$ cuts $D'D'$. A will plan to buy OF nuts at a price of OG per nut:

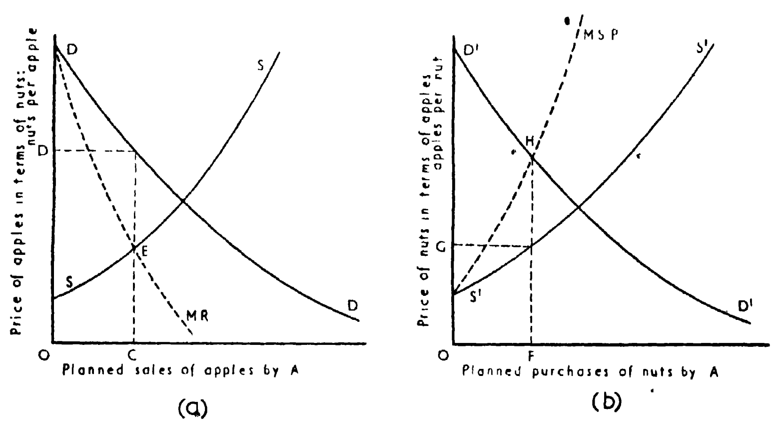


Diagram 121

OG is equal to UZ/UL_a in Diagram 120 and to $1/OD$ in Diagram 121 (q), and OF is the same as OJ (or ZN) in Diagram 120.

Third, we shall assume that both A and B try to be price-makers: that is, that A acts as a monopolist for apples and as a monopsonist for nuts, and that B acts as a monopolist for nuts and as a monopsonist for apples, or that A behaves conjecturally believing that B will act autonomously, and that B behaves conjecturally believing that A will act autonomously. The consequences of this assumption are portrayed in Diagram 122. ZR' and ZS' are the offer curves of A and B respectively. Knowing B's offer curve, A will plan to exchange the quantities implicit in L_a (where ZS' touches one of A's indifference curves) by fixing the rate of exchange at UL_a/UZ; knowing A's offer curve, B will plan to exchange the quantities implicit in L_b (at which ZR' is tangential to one of B's indifference curves) by fixing the rate of exchange at $U'L_b/U'Z$. In this model, it is clear that no equilibrium will be reached. The hypothesis that A (or B) makes about B's (or A's) behaviour will be proved wrong as soon as it is tested, and the assumptions by which our model is defined do not help us to identify what new hypothesis each party to the barter exchange will choose. The analogy between this model and the model of duopoly in which each firm aspires to output-leadership (see *supra*, pages 331–3) is apparent.

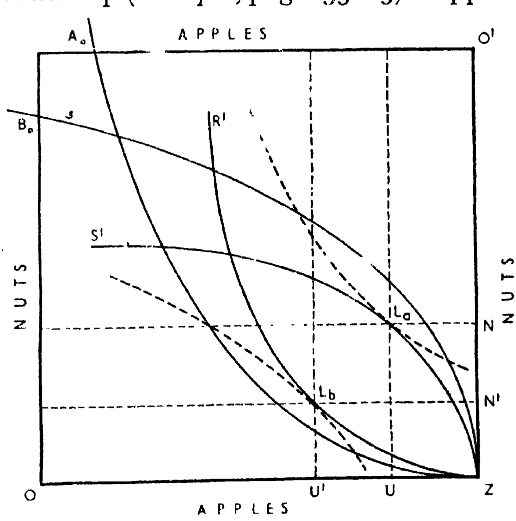


DIAGRAM 122

The information contained in Diagram 122 may be represented with the aid of demand and supply curves and their derivatives. In Diagram 123, the *DD*- and *SS*-curves have the same meaning as in Diagram 121 (a). The expected marginal revenue curve of *A*, the monopolist for apples, is shown by *MR*; we shall call the *MSP*-curve,* which is the marginal curve corresponding to *SS*, the marginal supply price curve of *B*, the monopsonist for apples. The monopolist *A* will plan to sell *OD* apples at a price of *OC* nuts per apple; *OD* is equal to UZ in Diagram 122, and *OC* to UL_a/UZ. The monopsonist *B* will plan to buy *OF* apples at a price of *OE* nuts per apple; *OF* is the same as $U'Z$ in Diagram 122 and *OE* is equal to $U'L_b/U'Z$. We can say no more than that the rate of exchange will lie somewhere between *OE* and *OC*, and that the quantity of apples exchanged for nuts will lie between *OF* and *OD*.

Lastly, we shall assume that *A* and *B* have decided to agree upon a rate of exchange by negotiating with one another. Let us suppose that each party knows the indifference map of the other, so that each is aware that the quantities exchanged must be denoted by a point lying within the area bounded by ZA_o

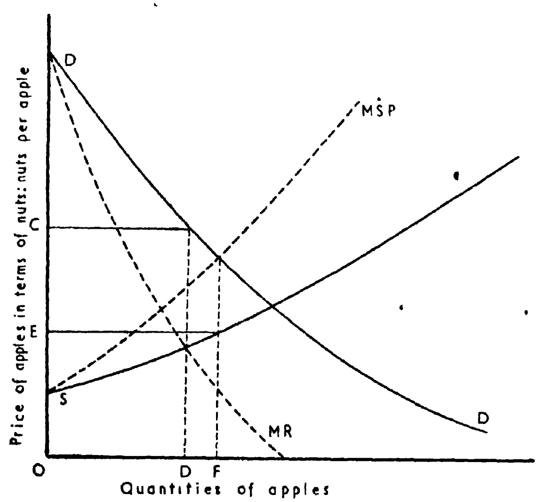


DIAGRAM 123

* For a description of the *MSP*-curve (though in a different context) see *supra*, page 305.

and ZB_0 in Diagram 124. Let us further suppose that A opens the negotiations by offering to sell B LZ apples in return for LP_1 nuts. It will be clear to both A and B that each of them can increase his satisfaction by moving north-westwards from P_1 within the envelope created by the indifference curves that cut one another at P_1. If they move to P_2, for example, each will still be able to enhance his satisfaction by moving north-westwards within the envelope created by the indifference curves that have one of their points of intersection at P_2. And so on, for if they start from any point near Z (Q), it will always pay A and B to increase (decrease) the quantities of apples and nuts that they exchange until some such point as C is reached, at which one of A's indifference curves is tangential to one of B's. The point C denotes a possible equilibrium position, for once it is reached, any movement away from it in any direction will reduce the satisfaction that is enjoyed by at least one of the parties. There will, however, be an infinite number of points such as C, and the one which is actually reached will depend on the point from which the negotiations begin and on the precise direction in which A and B move from it as the negotiations proceed. All these points will lie on the curve XY which is called the *contract* curve. If we start from any point on the contract curve, a movement away from the curve in any direction will reduce the satis-

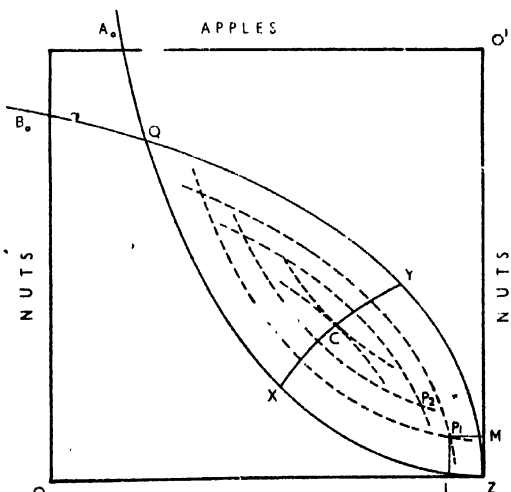


DIAGRAM 124

faction of both *A* and *B*, and a movement along *XY* will increase the satisfaction of *A* (or *B*) and reduce that of *B* (or *A*). Of all the solutions that lie on *XY*, *A* will prefer that denoted by *Y*, for that promises him the *maximum maximorum* of satisfaction, and he will not accept any solution lower than that shown by *X*; *B* will prefer that denoted by *X*, and he will not be willing to trade if the quantities of nuts and apples that are to be exchanged both fall short of the quantities denoted by *Y*. The contract curve is thus the locus of all possible outcomes of the negotiations and its length illustrates the range within which bargaining must take place.

We add little of substance to either our analyses or the conclusions we have drawn from them if we take other examples of bilateral monopoly. If we suppose, for example, that *A* is an employers' federation and *B* a trade union, we can develop a succession of models similar to those we have examined above, and we shall find that each yields the same kind of solution. If there is collective bargaining between *A* and *B* to help to determine the wage-rate, then, as in the previous paragraph, we can narrow the range of possible outcomes to those lying on some such line as *XY*; and the extent of the range will depend on the positions of *X* and *Y* — that is, the minimum acceptable solutions of *A* and *B* respectively — and these will depend on who or what *A* and *B* represent. We may say that the precise terms of the agreement will reflect the relative strengths of the negotiators, when 'relative strength' is defined in the same way as in the previous section, and we may proceed to list the things which help to determine a negotiator's strength as compared to that of his rival. The precise contents of this list will vary from one example of bilateral monopoly to another, but they will all be of the same kind as those listed in the concluding paragraphs of the preceding section.

USEFULNESS OF THE THEORIES OF MONOPOLY, OLIGOPOLY, ETC.

Throughout the first eight chapters, we assumed that each buyer and each seller was a price-taker. In Chapter 9, we described what the real world would have to look like if each household and firm in it was to behave in this way in the markets in which it bought and sold products and productive ser-

vices. That particular assumption about market behaviour, and the market morphology that was generally associated with it, was useful and realistic because it helped us to predict the consequences for relative prices of present economic events, and to explain past changes in the pattern of relative prices. In this and in the previous chapter, we have sketched the bare bones of a classification of the market morphologies that are to be found in a modern, free-enterprise economy, and we have described the kinds of market behaviour that are likely to be associated with each class: thus, a monopolist may be a price-taker or 'market-divider'; a monopolistic competitor may be a price-maker, advertiser or 'product-differentiator'; an oligopolist may act autonomously, conjecturally or collusively in fixing a value for his price (or output), product-quality, advertising or research expenditures. We shall conclude this chapter by discussing the usefulness of these models of 'impure and imperfect' competition.

We have argued in Chapter 9 that the test of the usefulness of an economic model is the empirical validity of the explanations and predictions which it offers. Our model of pure competition was judged useful because it helped us to explain and predict changes in the pattern of relative prices, and to understand how a price system determines what goods shall be produced, how, when, where and in what quantities they shall be produced, and who shall have the privilege of acquiring them. If we are still broadly concerned with the *direction* of changes in relative prices, the models of monopoly, monopolistic competition, oligopoly and bilateral monopoly, add little (if anything) to the explanatory and predictive power of the model of pure and perfect competition. Thus, in Chapter 5, we showed that if the demand for some good X should rise, in the short-run its relative price would tend to rise also. If X is produced and sold by a monopolist, the same consequences will generally follow: since the price at which a monopolist will maximise his net revenue is equal to $MC/(1 - 1/e)$ (see *supra*, Chapter 10, pages 283 and 285), the change in his price which follows a rise in demand will depend on the behaviour of marginal costs and on the price elasticities of the new demand curve. It follows that his price will be lower after the rise in demand in two circumstances only: first, if his marginal cost is falling and if the price elasticity of the new demand curve is not sufficiently less (at the new profit-

maximising price) than that of the old one (at the initial profit-maximising price) to offset the rate of fall in marginal costs;* and second, if the rise in the marginal costs is more than offset by the rise in the elasticity of demand.† It is unlikely, however, that either of these will be generally true. Similarly, if there is a permanent rise in the 'demand' for the group of products produced by a monopolistically competitive industry, the price of each firm's product will ultimately tend to rise; or if we are given the relative strengths of an employers' association and of a trade union, and if the demand curve for the product rises, this will generally cause a widening of the bargaining range, so that we would expect the wage-rate to rise also.

It would seem, then, that we can answer the question: 'In what direction will relative prices alter as a consequence of some present economic event?' by using any one of the models which we have so far examined in this book. Since the model in which each buyer and each seller is assumed to behave as a price-taker is the simplest to use, it is that which we would generally choose. The fact that this model is the most abstract of all that we have discussed should not surprise us, for the degree of abstraction in our concepts varies directly with the generality of the problems we hope to elucidate with their aid. As we demand specific answers to less general questions, however, we must rely increasingly on the models of monopoly, oligopoly, etc. We shall give some examples of the kinds of questions to which these models give answers that conform more closely to the facts than those provided by the model of pure and perfect competition.

First, suppose we are told that (1) there was a permanent rise in the demand for commodity *X*‡ on 1/1/56; (2) the price of *X* had not altered by 1/4/56, and (3) the prices of the productive services needed to produce *X* did not alter, nor were they expected to alter, and that we are asked to explain (2). Our analyses offers at least the following explanations:

(i) *X* is produced by a perfectly competitive industry in which

* Thus, if marginal costs fall from 10 to 8, and if elasticity falls from 2 to 1·8, then price will fall from 20 (that is, $10/(1-\frac{1}{2})$) to 18 (that is, $8/(1-1/1{\cdot}8)$).

† Thus, if marginal costs rise from 10 to 12, and elasticity rises from 2 to 4, then the price will fall from 20 (that is $10/(1-\frac{1}{2})$) to 16 (that is, $12/(1-\frac{1}{4})$).

‡ Or a group of commodities which are close substitutes for one another. If we think of *X* as a product-group, the word industry hereafter must be interpreted as 'industry'.

the calendar time required for long-run adjustments to be effected is less than three months;

(ii) X is produced by a monopolist, and either (*a*) the new demand curve for X was iso-elastic with the old one and marginal costs were constant over the relevant range of output; or (*b*) he reduced his costs by altering his technique of production; or (*c*) as the demand for X rose, the monopolist began to fear that new firms might be set up to compete with him, so that he maintained his price to deter them;

(iii) X is produced by an oligopolistic 'industry', and either (*a*) there is price-leadership and either (ii) (*a*), (*b*) or (*c*) above is true; or (*b*) prices are fixed by an informal agreement that is enforced by the hypothesis on which the kinked demand curve model was based, and each firm's marginal cost curve cuts the new marginal revenue curve within its 'discontinuity'; or (*c*) the oligopolists have agreed to pursue the monopoly solution, and are either deterred from raising the price by (ii) (*c*) above or are awaiting the expiration of their existing agreement.

We can decide which of these groups of explanations is likely to be relevant by counting the number of firms that produce X. If we should find that the number is small, so that the explanation must lie within group (iii), we must seek further information about the industry (or 'industry') to decide between (iii) (*a*), (*b*) and (*c*). Once we have discovered the proximate cause of the price stability, we can list the other changes which are likely to have been associated with it, and which are likely to occur in the future. Thus, if we find that X is produced by an oligopolistic industry (or 'industry') in which there was an explicit or implicit price agreement, we would expect each firm to be attempting to enhance its sales (at the expense of its rivals) by placing greater emphasis on advertising, variations in product-quality and research; and we might expect also that in future some or all of the firms will demand that the existing agreement be revised. If the existing agreement promised the monopoly solution in the light of the old demand conditions, the participants are likely to have become increasingly dissatisfied with it, so that it will probably be presently dishonoured or formally terminated and new negotiations initiated.

Second, suppose we are told that there has been a permanent

rise in the demand for some commodity (or some class of commodity) X, and that we are asked to predict the probable consequences of the introduction of price controls which fix the legal maximum price of X at its old level. Our predictions will depend on the morphology of the market in which X is produced:

(i) If X is produced under conditions of pure and perfect competition, then (*a*) there will be an 'excess demand' for X at the legal maximum price; (*b*) the previous pattern of allocation of X amongst households will be upset, for only those who shop earliest will be able to obtain X; (*c*) if the government desires an 'equitable' distribution of X, it must introduce its own rationing scheme, and (*d*) even with official rationing, the legal maximum price will be difficult to enforce and a black market is likely to appear.

(ii) If X is produced by a monopolist, then (*a*) the enforcement of the legal maximum price will be less difficult for it will be easier to control the total supply of X when there is only one producer and seller to police; (*b*) the monopolist is the less likely to evade the controls or condone or patronise a black market for he can be identified as the culprit not only by the government but also by his clients, and he may be loathe to sacrifice their goodwill; (*c*) the monopolist may increase his output at the controlled price: thus, in Diagram 125, AR and MR are the new

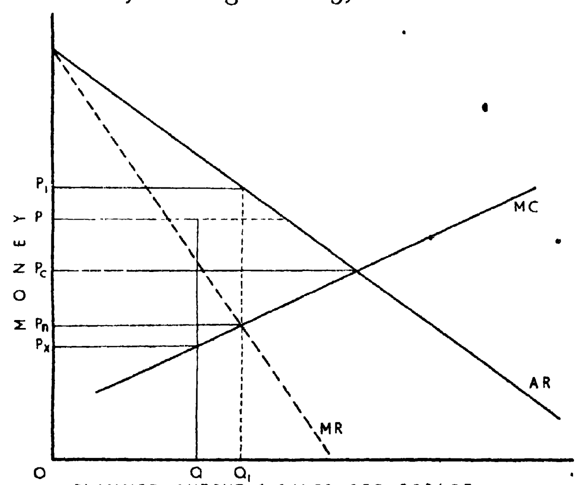


DIAGRAM 125

demand and marginal revenue curves, OP and OQ the price which he charged and the output he produced before demand rose, and OP_1 and OQ_1 are the price and output which he would choose now to maximise his net revenue if no price control had been introduced; it is clear that if OP is less than OP_1 (as it must be) and greater than OP_x,* the monopolist will increase his output above its pre-control level,† and that at prices between OP_1 and OP_n he will produce an output greater than that which he would have produced had there been no price control; (*d*) if there is an excess demand for X at the controlled price (that is, if OP lies between OP_c and OP_x), official rationing may be unnecessary, for the monopolist will be loathe to destroy the goodwill of his clients by favouritism or inequitable distribution; (*e*) the monopolist may be tempted to enhance his net revenue by reducing the quality of his product, though the effects of this on the demand for his product after price control has been removed may help him to withstand this temptation.

(iii) If X is produced by a monopolistically competitive 'industry' the probable consequences of price control will be similar to those that we have listed above for monopoly.

(iv) If X is produced by an oligopolistic industry (or 'industry'), then (*a*) if the products are differentiated, we would expect the same range of probable consequences as we have listed under (ii) above for monopoly; (*b*) if implicit or explicit price agreements have been common in the industry, the task of the price control authority will be the easier, for a method of price control and a mechanism for its enforcement will be available: thus, if there is price leadership, the controlling authority need do little more than fix the leader's price; (*c*) whether or not price agreements have been common, price control merely removes one parameter of action from the armoury of each firm, so that each may continue to strive to increase his share of the market by varying his advertisement, quality and research; (*d*) even if rising production and selling costs drastically reduce the net revenue of each firm, there may be no concerted attempt to

* We shall suppose for simplicity's sake that by selling OQ at OP_x per unit, the monopolist will cover his total costs of production.

† For at any given controlled price, the monopolist will maximise his net revenue per period by producing the output at which his marginal cost is equal to that price. In Diagram 125, there will be an excess supply at prices above OP_c, and there will be an excess demand at prices between OP_c and OP_x.

have the legal maximum price raised, for each firm might feel that it can establish a permanent lien on a larger share of the market by accepting the existing maximum price while the control continues.

(v) If there is bilateral monopoly in the market in which X is bought and sold, there is little that we can say. The circumstances which caused the introduction of price control will certainly affect the bargaining limits and may also cause changes in relative strengths. We may, if we like, view the advent of the price control authority as creating 'trilateral monopoly', for it is likely that it will seek to protect the 'community' or the 'public interest' or 'consumers' — that is, a party whose interests have not customarily been either pleaded or heeded.

While the above list of the possible consequences of price control is by no means exhaustive, it suffices to show that the effects may vary from one market morphology to another, so that we cannot be indifferent as to which model we use. There are two reasons why the models which we have examined in this and in the previous chapter yield different predictions from the model of pure competition. First, the smaller is the number of sellers, the greater the likelihood that each will develop a sense of 'property' in the market for his product: if there are very many sellers of a homogeneous product, then black market dealings, profiteering and concealed reductions in quality may be common, for it is not easy for consumers (or the government) to find the culprit; if there is only one seller, this is not so, and in addition, he may reasonably expect that his future reputation and profits depend on his present actions.* Second, the number of variables over whose values each firm has control, and the limits within which it may fix their values, vary from one market-form to another. In pure competition, quantity of output is the sole parameter of action of each firm; in differentiated and non-collusive oligopoly, the parameters may include price/output, advertisement, product-quality and research expenditures. We would expect the range of possible consequences to vary more or less directly with the number of variables whose values lie within each firm's control. These same reasons offer a partial

* This reason is roughly the same reason why there is no litter in my garden and much litter in the public park.

explanation of why the task of the price control authority may be relatively easy in some markets and more difficult in others.

Finally, we shall list briefly two further uses of the models of 'impure and imperfect' competition. First, they are toys: toys help children to acquire notions of distance, texture, form, mass, weight, and perhaps even of property, and to develop manipulative skill. In these ways, they help to prepare children for the real world in which they must live and work. All the models that we have discussed in this book serve a similar purpose. Only by using them will a student become aware that the economic world is a world of relationships and not of independent quantities. The language of economics is mainly that of everyday speech, and the words like demand, cost, output which mean relationships in economics mean simple quantities in everyday usage; geometric models acquaint the student with these relationships, and with the circumstances in which they will alter, and when he is so acquainted, he is more than halfway to being an economist. Further, by manipulating these models, a student develops a 'feel' for interdependence and mutual causation, and is in that way properly conditioned for his own explorations of the more complex economic reality. Second, these models of 'impure and imperfect' competition may assist the student of welfare economics. If we wish to pass judgement on actual markets or industries, the first step is to discover whether they deviate from our 'ideal', and if they do, to measure the extent of the deviation and to distinguish its proximate causes. If the model of pure and perfect competition is chosen as the 'ideal', then by comparing it with the models of monopoly, monopolistic competition, oligopoly and bilateral monopoly, we can roughly accomplish these tasks.

CHAPTER 12

Some Further Problems

The analysis in the previous chapters provides an introduction to the theory of relative prices. In the model developed in the first eight chapters, we assumed broadly that each actual and potential firm made (or might make) only one product, and that the market in which it was sold (or might be sold) was both purely and perfectly competitive. In Chapter 9, we justified the use of this model by showing that it gave tolerable answers to many broad and general questions about the behaviour of relative prices. In Chapters 10 and 11, we showed how imperfect competition (in the sense in which we defined it) can provide a first approximation to a theory of profit, and we modified some of the assumptions of pure competition (mainly those relating to the numbers of buyers and sellers) and catalogued some of the consequent modifications in our explanation of the determination and behaviour of relative prices. Thus far in this book, we have ignored the influence of the organisation of a firm on its aims and behaviour; we have not dealt with multi-product firms; we have virtually ignored time and the related (though not identical) problems posed by imperfect knowledge and uncertainty, and we have assumed throughout that the objective of the firm is to maximise its net revenue. We have not described the actual rules, institutional arrangements and market practices by which prices are fixed and by which competition is curbed. Finally, we have relied mainly on the neo-classical tools of analysis — namely, revenue and cost functions and their derivatives.

In this, the final, chapter we shall do three things. First, we shall discuss the choice of a sales plan by a multi-product firm; second, we shall examine critically the actual method by which firms appear to fix the selling prices of their products; and third, we shall describe briefly two alternative techniques of analysis

— namely, mathematical programming and the theory of games.*

THE MULTI-PRODUCT FIRM

A firm may produce more than one product either from necessity or from choice. Where the necessity is imposed by technical conditions or natural causes, we generally speak of *joint production*, and typical examples are beef and hides, wool and mutton, petrol and oil, meat and offal. Alternatively, a firm may decide to produce several products, because it is profitable for it to do so. Thus, the production and sale of product *A* may lower the costs of producing and/or selling *B*.

The proportions in which joint products appear may be fixed or variable. If two or more products can be produced only in fixed proportions — for example, if the manufacture of one unit of *A* always yields two units of *B* — our analysis in Chapter 2 requires no modification, for we can treat one *A* plus two *B* as constituting a single unit of a new product, *X*. The firm can calculate the total costs of producing each possible output of *X*, and so obtain a total cost schedule. If the firm is a price-taker in the markets in which *A* and *B* are sold, it can estimate the total receipts that would be obtained by selling each output of *X*, and thus obtain a total revenue schedule. From these schedules the firm can decide what output and sales plans to implement for *X*, in the light of its objective.

This sales plan will be revised if there is any change in the expected selling prices of *A* or *B*, or in the costs of producing them, or in the firm's objective. Thus, if the price of *A* rises, the price of *B* remaining constant, the planned production and sales of *A* and *B* will rise and by the same amount. By tracing the behaviour of the planned sales of *A* (or *B*) as the price of *A* (or *B*) is varied, we can obtain a supply curve for *A* (or *B*). This supply curve will have the same shape as the short-run supply curves described in Chapter 2 above, and for the same reasons.

The only difference is that the position of the supply curve for *A* (or *B*) will depend on the price of *B* (or *A*), for more of *A* (or *B*) will be offered for sale at any given price the higher is the price of *B* (or *A*) and conversely.*

The argument is similar, though more complicated, if the joint products can be produced in variable proportions. Here, as before, the firm can calculate the minimum costs that it would incur in producing each combination of the two (or more) products, and, given their prices or demands, it can estimate the revenue that the sale of each combination would yield. From these cost and revenue data it will choose the production and sales plans which promise to fulfil its objective.

We can illustrate this choice on diagrams similar to those used in Chapter 2, and by doing so we can see more clearly the patterns inherent in the data which inform and limit the firm's choice. In Diagram 126, we measure quantities of product *A* on the horizontal axis, and quantities of product *B* on the vertical axis. Between these axes, each combination of *A* and *B* that the firm might produce can be plotted, and beside each can be written the total costs that would be incurred in producing it and the total revenue that would be obtained from selling it. We can get a picture of the patterns implicit in the cost and revenue

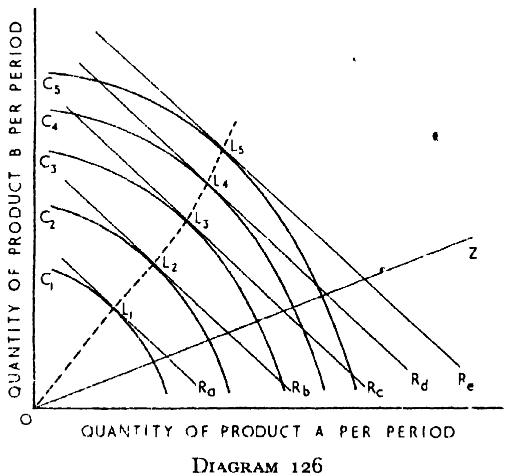


DIAGRAM 126

* The description of the choice (or revision) of a sales plan by a monopolist who is producing joint products in fixed proportions is formally similar and is left to the reader.

data by drawing cost contours,* each of which passes through all combinations of A and B that are equally costly to produce, and revenue contours, each passing through all combinations which promise the same receipts. In Diagram 126, the former are labelled $C_1, C_2, \ldots C_n$, and the latter $R_a, R_b, \ldots R_m$. We shall first explain the shapes of these contours and then their positions relative to one another.

We assume that each cost contour is concave when viewed from the origin — that is, as the proportion of B (or A) to A (or B) rises, it becomes more and more expensive to effect further substitution of B (or A) for A (or B). The fact that the revenue contours are straight lines reflects our assumption that the firm is a price-taker in the markets in which A and B are sold; if the firm had been to any degree a price-maker in either or both of these markets, then the revenue contours would have been more or less convex (or a mixture of the concave and the convex) when viewed from the origin. If we suppose that A is mutton and B wool, and that each breed of sheep yields mutton and wool in fixed proportions,† then any straight line OZ radiating from the origin will summarise the behaviour of costs as the output of breed Z is increased. We would expect the distances between successive cost contours as we move along this line to become progressively smaller, because of the operation of the law of diminishing returns. If the firm is a price-taker in both the mutton and wool markets, the distances between successive revenue contours will be equal to one another, for revenue will vary directly with sales. If the firm is a price-maker in either or both markets, these distances will depend on the behaviour of the elasticities of demand for mutton and wool as their prices fall with expanding sales.‡

Since we suppose that the objective of the firm is to maximise its net revenue, then, of all the combinations of A and B lying

* We avoid the term 'iso-cost curve' here, for that was used to describe a different curve in Chapter 2.

† The fact that each cost contour is a smooth curve reflects our implicit assumption that there is an infinite number of breeds of sheep.

‡ If the demand for each commodity has unit elasticity throughout its length, then the revenue contours will coincide with one another. If the demand curves for mutton and wool are curvilinear with varying elasticities, there is little that can be said about the relative positions of the revenue contours, except that they will for a while assume progressively larger numerical values and thereafter progressively lower numerical values.

on C_1 which might be produced for (say) £100, it will choose that which yields the maximum revenue — that is, that denoted by the point L_1 at which C_1 is tangential to a revenue contour. In choosing its sales plan, therefore, the firm will confine its choice to the range of plans lying on the line $OL_1L_2L_3 \ldots$, and of these eligible plans it will decide upon that which promises the maximum net revenue per period.

The manner in which the firm's sales plan will be revised in response to a change in the expected relative selling prices of A and B can be deduced from Diagram 126. If the price of B rises as compared with that of A, then each revenue contour will rise less steeply, and we would expect the line equivalent to $OL_1L_2L_3 \ldots$ to lie to the left of its previous position. The firm would, therefore, now plan to produce and sell more of B, and less of A. From the successive revisions of the sales plan in response to increases in the price of B (A), the price of A (B) remaining constant, we can obtain a supply curve for B (A); its position will clearly depend on the level of the price of A (B).*

In this section, the choice of a sales plan by a multiple product firm has been couched in terms of the relationship between total cost and total revenue. We might now proceed to translate our conclusions into terms of relationships between marginal costs and revenues, as we did in Chapter 2. There is, however, little to be said for attempting to do so. Where the joint products can only be produced in fixed proportions to one another, it is not possible to distribute the additional costs incurred in producing 1A plus 2B between A and B, so that the term 'the marginal cost of A (or B)' has no meaning. Where the proportions are variable, the marginal costs of producing any additional unit of A (or of B) can be isolated, but there is no reward for isolating them since all that emerges is a translation of our old conclusion and not a new one.

* The analysis of the case in which the firm produces multiple products by economic choice rather than technical necessity is formally the same as that illustrated in Diagram 126. If an increase in the production of A (the output of B remaining constant) reduces the costs which can be unambiguously attributed to the production of a given output of B, then the cost contours will be 'flatter' than in Diagram 126: that is, that part of C_1 near the vertical axis will tend to lie south of its present position and the range of C_1 as it curves towards the horizontal axis will lie to the east of its present position.

FULL- OR AVERAGE-COST PRICING*

The method of analysing relative prices which we have used is called 'marginalism' or the 'marginal analysis'. It is so called because marginal cost and marginal revenue curves are generally used in the diagrams which illustrate the firm's choices. These functions are derived from the total cost and total revenue curves respectively; they are probably unnecessary, for all the firm's choices can be illustrated with the total curves themselves. Indeed, the total curves are preferable for they clearly illustrate the different combinations of cost and revenue open to a firm in a defined situation, and thus describe the range from which it must choose its sales or purchase plans. If marginal curves are used, there is the danger than an economist (or his audience) might suppose that a firm thinks marginally — that it consciously seeks the sales plan which promises to fulfil its objective by experimenting with increments and decrements in its output and sales.

The full-cost or average-cost theory of price purports to be a description of how the typical businessman actually fixes the selling price of his product. This theory usually rests on statements by businessmen or on questionnaires which they have completed. It may be summarised as follows: †

(i) 'the price which a business will normally quote for a particular product will equal the estimated average direct costs of production plus a costing margin.' It is assumed that the average direct cost function will tend to be a horizontal straight line over a part of its length if the prices of the direct cost factors are given.

(ii) 'the costing-margin will normally tend to cover the costs of the indirect factors of production and provide a normal level of net profit, looking at the industry as a

whole.'. Once chosen, the costing-margin will remain constant, 'given the organisation of the individual business, whatever the level of its output.' It will tend to vary, however, with 'any general permanent changes in the prices of the indirect factors of production'.

(iii) 'given the prices of the direct factors of production, price will tend to remain unchanged, whatever the level of output.'

(iv) 'at that price, the business will have a more or less clearly defined market and will sell the amount which its customers demand from it.'

This method of fixing prices, which is said to be followed by all or most price-makers, is illustrated in Diagram 127. In accordance with the assumption in (i) above, the average direct cost curve (and therefore the marginal cost curve) is a horizontal straight line over a part of its length. If we equate the indirect costs with what we have called fixed costs, then they, together with the profit which the firm expects, hopes or plans to earn, will give a fixed sum of money which will be a datum for any existing firm making a short-run plan. The absolute amount of the costing-margin is derived from this sum of money by dividing it by some output. This output might be determined either (*a*) as a percentage of capacity output, where capacity is inter-

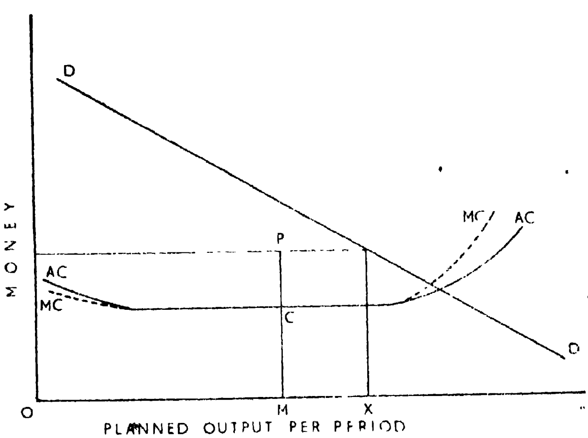


DIAGRAM 127

preted as an engineering fact; (*b*) as the output which was sold in the preceding production period, or the average of realised sales over a number of past production periods; or (*c*) as the minimum, mean, median or modal output that the businessman expects to be able to sell in a future period. If the firm is a new one, or if it is an existing firm introducing a new product, then only the first and third of these interpretations will be relevant; in these circumstances, indeed, it is likely that the first will coincide roughly with the third, for the capacity of the plant will depend on expected future sales. We shall assume that the firm chooses the output *OM* as the basis for its choice. Its selling price will therefore be equal to *MC* plus the costing-margin *PC* — that is, to *MP*. If *DD* is the demand curve for the firm's product, then at the price *MP* per unit the firm will succeed in selling *OX* per period. This price will not be altered in response to changes in demand, but only in response to changes in the prices of the direct and indirect factors.

One difference between the average-cost theory and that in the preceding chapters lies in the shape of the average direct (or variable) cost curve. The shape given to the average cost curve in Diagram 127 is in general accord with the results of empirical investigations into cost behaviour.* The shape which we have generally given to the average variable cost curve was explained by the law of variable proportions and by our assumption that the businessman when making a long-run plan had in mind some precise output which he will decide to produce with the bundle of 'fixed' and other factors that promises him the maximum net revenue. If an entrepreneur expects his sales per period to vary widely, however, it may be profitable for him to choose factors which limit the rate at which physical returns diminish and thus promise virtually constant average direct costs over a wide range of outputs.† This difference, however, is not crucial, for our previous analysis and conclusions would need no substantial revision were all the average variable cost curves given flat bottoms.

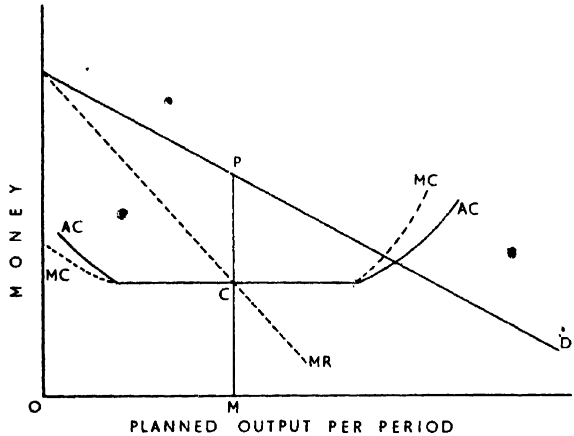


DIAGRAM 128

In Diagram 128, we compare the average-cost theory with that developed in this book. If the firm is in any degree a price-maker, it will (according to our analysis) plan to sell *OM* per period at a price of *MP* per unit. We may say that the firm chooses this price by taking its average direct costs of production and adding a costing-margin. In the marginal analysis, the size of the costing-margin, *PC*, depends *inter alia* on our assumptions that the firm knows (or thinks it knows) the costs and demand for its product and that it seeks to earn the maximum net revenue from its operations. In the average-cost analysis, the costing-margin may be *PC*, or it may be adjusted to approximate towards *PC*.* If the costing-margin suggested by the average-cost theory differs from *PC*, the explanation might lie in the fact that the businessman (*a*) is quite unaware of his costs and demand, or (*b*) pursues some maximand other than money profit. A firm is unlikely to be completely ignorant of the demand for its product, for at worst it can ascertain something about the current demand for products that are close substitutes. A multi-product firm may find it too costly both in terms of time and of money to proceed in the manner described in the first section of this chapter; it may, therefore, take that part of the costs of production which can be unambiguously attributed to any product *X*, and determine the price of *X* by adding a margin. If the firm is trying to maximise its net revenue, however, we

would expect that the addition of this margin would give a price which would approximate towards that suggested by our analysis of the multi-product firm. It would appear, therefore, that if the two theories suggest different prices, the cause must lie in firms which follow the average-cost theory pursuing some objective other than maximum money profits. There seems to be little evidence that businessmen do not seek to earn the greatest net revenues: when they have other aims, such as security, the extent to which these are achieved will usually vary more or less directly with the profits earned; where they appear to eschew the quest for the maximum profits, the explanation is likely to lie in idiosyncracies in their expectations about the behaviour of demand and cost or in the extent of their planning horizon.

If we accept the criticisms of the previous paragraph, then the average-cost theory merely describes the method (or the initial stages in the method) by which firms seek to fix their prices at the levels suggested by our previous analysis. While following the practice, they seek the maximum net revenue by varying the size of the costing-margin in the light of their expectations of demand. The manner in which the average-cost theory is frequently stated, however, implies that the costing-margin is invariant when demand alters. This is probably untrue, but if it were a true description of business practice than our previous analysis would need to be substantially modified, for it would mean that the price of a product responded solely to changes in the prices of the direct and indirect factors used in producing it and that it did not respond in any way to changes in the demand for it.*

MATHEMATICAL PROGRAMMING†

In Chapter 2, we distinguished between the different methods or techniques by which a genus of product might be produced. A method of production was defined by two criteria: first, the species of the product, and second, the kinds and qualities of the

factors of production that were used, varied from one method to another. We assumed that within any method of production, the rate of output could be varied by changing the relative quantities of the relevant factors — that is, by altering the quantities and proportions in which they were employed. In Chapter 3, we assumed, *inter alia*, that there was an infinite number of methods of producing a product, and that the proportions in which the appropriate factors could be employed were continuously variable. On these assumptions, we illustrated the firm's choice both of a method of production and of the extent to which it should be utilised to achieve the firm's objective.

In mathematical programming or activity analysis, a basic concept is the 'process' or 'activity'. A process is 'a specific method of performing an economic task'. For example, 'the weaving of a specific quality of cotton grey goods on a specific type of loom' is a process. The long-run plan of a firm is simply its decision 'as to which processes are to be used and the extent to which each is to be employed'.* The notion of a process or activity is, therefore, closely akin to that of a method of production. In mathematical programming, however, it is generally (though not necessarily) assumed (*a*) that the number of processes is finite, and (*b*) that within a process, the relevant factors are used in fixed proportions and output varies directly with input.

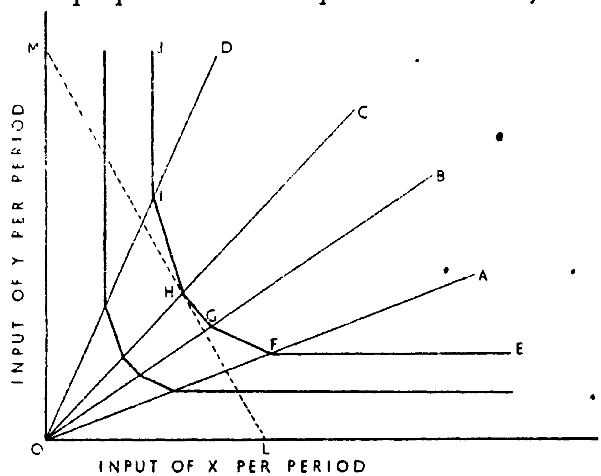


DIAGRAM 129

* All quotations in this paragraph are taken from Dorfman, *loc. cit.*, page 798.

A firm's choice of a process and of the extent to which it should be used are illustrated in Diagram 129.* We assume that each process by which the same class of product might be produced yields precisely the same species of it. Further, we assume either (*a*) that all processes use precisely the same kinds and qualities of factors, but that the proportions in which they are fixed varies from one process to another, or (*b*) that while the qualities of the factors varies from one process to another (for example, different processes use different grades of labour), these different qualities can be reduced to common physical units (for example, homogeneous man-hours per period). We suppose that the production of product *Z* requires two factors, *X* and *Y*, and that *Z* can be produced by four processes, *A*, *B*, *C* and *D*. In the diagram, the straight line *OA* shows the proportion in which *X* and *Y* must be combined in process *A*. Beside each point on *OA*, we can write the output of *Z* which that combination of *X* and *Y* would promise. If the inputs of *X* and *Y* can be varied by infinitely small increments then a continuous range of outputs of *Z* can be obtained from *A*; if the inputs of *X* and *Y* can be increased only by finite increments then there will be a finite number of possible outputs and the points denoting them on *OA* will be equidistant from one another on *OA* if there are constant returns in terms of *Z* to inputs of *X* and *Y*. We shall suppose that the output of *Z* promised by the combination of *X* and *Y* denoted by *F* on *OA* is the same as that promised by *G* on *OB*, *H* on *OC* and *I* on *OD*. If the processes can be continuously substituted for one another, then the discontinuous line *EFGHIJ* will pass through all combinations of the processes which promise the same output of *Z*;† if no substitution is possible between processes, then this particular output of *Z* can only be produced by the four combinations of *X* and *Y* which are denoted by *F*, *G*, *H* and *I* respectively. The line *EFGHIJ* is analogous to an isoquant. On our assumptions, each segment of an 'isoquant' in Diagram 129 will be parallel to the segment between the same process lines, and the 'isoquants' will be equidistant from one another.

When the prices of *X* and *Y* are given, price lines can be

* Diagram 129 is based directly on Figure 5 in Dorfman, *loc. cit.*, page 807.

† For the proof of this proposition, see Dorfman, *loc. cit.*, Figure 4, page 805 and the text on page 806.

drawn: thus, the straight line *LM* passes through all combinations of *X* and *Y* that could be purchased for some given sum of money. If the firm were planning to expend this sum, it would buy the quantities of *X* and *Y* denoted by *H* and combine them in process *C*, for only by doing so would it obtain the maximum output of *Z*. On our assumptions, if the prices of *X* and *Y* remain constant, the combination of *X* and *Y* which promises the maximum output for any given expenditure will always lie on *OC*, so that *OC* will be analogous to an expansion path. From the information that is implicit in this 'expansion path' the firm can obtain relationships between total costs and total revenue from which it can choose its sales plan (that is, the degree of utilisation of process *C*) in the light of its objective. If the relative prices of *X* and *Y* are such that the price line coincides with a segment (such as *GH*) of the 'isoquant' then the firm's aims might be achieved by using process *B* to *G*, or process *C* to *H*, or by any one of the combinations of *B* and *C* lying on *GH* if the two processes can be continuously substituted for one another.

Diagram 129 illustrates the firm's choice of long-run sales and purchase plans — that is, the same decisions as we analysed in Chapter 3. The fact that Diagram 129 closely resembles Diagram 27 by which we illustrated a firm's short-run decision is solely a consequence of the assumptions on which it rests: if the factors must be used in fixed proportions in each process then each isoquant contracts to a point and all such points will lie on a straight line such as *OA*; if output varies in the same proportion as input within each process then there will be constant returns to scale, so that the 'isoquants' will be 'parallel' to one another.

Thus far we have assumed that the firm can buy unlimited quantities of *X* and *Y* at constant prices. In mathematical programming, more importance is attached to the case in which one or more factors can be bought at given prices only up to a fixed maximum quantity. The effect of a scarce factor is shown in Diagram 130, in which the process lines and 'isoquants' of Diagram 129 are reproduced. We shall suppose that the firm cannot acquire more than *OK* of factor *Y*: thus, *Y* might be floor space and *OK* might be the firm's present factory space which cannot be extended, or *Y* might be money with *OK* representing the maximum sum that the firm can spend on equip-

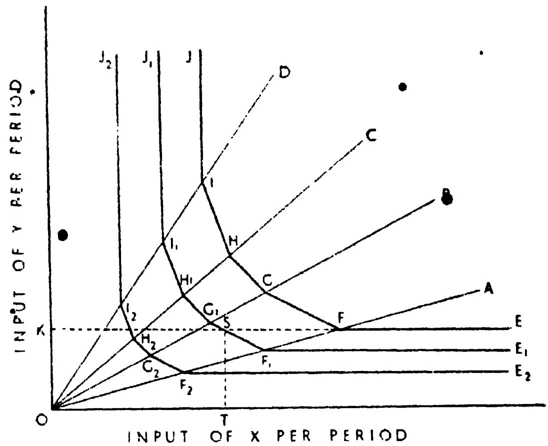


DIAGRAM 130

ment. In these circumstances, the firm will be confined to the production possibilities lying on KE, for it will only be able to increase its output of Z by using more of X in conjunction with OK of Y; from this range the firm will choose that combination of factors and processes which promises to fulfil its objective. If we assume not only that the firm can acquire no more than OK of Y, but that it can hire no more than OT of X, then S will denote the highest output that the firm can produce and this output will also be that which promises the maximum net revenue, where together with our other assumptions the prices of X and Y remain constant up to purchases of OK and OT respectively, and where the firm is a price-taker in the market in which it sells its product.

In this section, we have applied mathematical programming only to the theory of production, and in doing so it has been reduced to its simplest form. The distinguishing characteristics of activity (as compared with marginal) analysis are first, the assumption that factor-proportions are fixed and returns to scale constant within each method of production; second, the assumption that factors can only be substituted for one another by substituting processes; and third, the importance attached to scarce factors. If we think merely in terms of the simple examples which we have taken, then the choice between the two types of analysis must depend on the empirical validity of these assumptions and of the conclusions which they yield. To so confine our-

selves, however, would accord much less than justice to mathematical programming. The main advantage of the definition of process, as compared with our definition of method, is that it is empirically recognisable, and the relationship between inputs and output within a process can be estimated with comparative ease. The analysis of the effects of scarce factors enjoys the same advantage as compared with our short-run analysis in Chapter 2. This close relationship between the actual structure of production and the concepts of activity analysis, together with the existence of practicable mathematical techniques, means that many actual problems in the fields of production and distribution can be solved numerically. So far, mathematical programming would appear to have been most successfully applied to the allocation and use of resources within a single firm. There can be little doubt that it will ultimately be applied to the wider problems of resource allocation and use within broad sectors of the economy or to the economy as a whole.*

THE THEORY OF GAMES†

The present position of the theory of games in the analysis of relative prices is similar to that of mathematical programming. While both of these techniques seem to offer little more than a restatement of some of the simpler analyses, they hold out the prospect of solutions to more complex problems which without them would be intractable or insoluble. We shall first attempt to translate our analysis of bilateral monopoly in Chapter 11 into the language of the theory of games. We shall suppose that *A* has a basket of apples and *B* a basket of nuts, and that *A* and *B* have different tastes and preferences. In Chapter 11, we described merely three kinds of behaviour or 'strategy' by which *A* (or *B*) might attempt to maximise his satisfactions, and we

showed that (as in duopoly) the success that attended A's (or B's) efforts depended on B's (or A's) behaviour. If we assume that the satisfactions enjoyed by both parties can be cardinally measured in common units (which we shall call 'utils'), and if we let the suffixes 1, 2 and 3 denote particular strategies,* then

TABLE 3

B's Strategies / *A's Strategies*	B_1	B_2	B_3
A_1	5	3	1
A_2	6	4	2
A_3	9	8	7

(*a*) A's satisfactions.

B's Strategies / *A's Strategies*	B_1	B_2	B_3
A_1	8	10	18
A_2	7	6	11
A_3	3	5	10

(*b*) B's satisfactions.

TABLE 4

B's Strategies / *A's Strategies*	B_1	B_2	B_3
A_1	5	3	1
A_2	6	4	2
A_3	9	8	7

(*a*) A's satisfactions.

B's Strategies / *A's Strategies*	B_1	B_2	B_3
A_1	10	12	14
A_2	9	11	13
A_3	6	7	8

(*b*) B's satisfactions.

* For example, the suffix 1 might denote autonomous behaviour with zero conjectural variation — i.e., A (B) will not change the rate at which he is willing to sell his apples (or nuts) if there is any change in B's (A's) offer of nuts (apples). If both A and B follow this strategy, then the outcome will be that illustrated by V in Diagram 118. Suffix 2 might denote autonomous behaviour with conjectural variation — that is, A (or B) will change his exchange rate in some constant (absolute or percentage) way if B's (A's) offer of nuts (apples) alters. If both follow this strategy, then an equilibrium similar to that shown by V in Diagram 118 will be reached. Suffix 3 might denote the quest for leadership. Thus, if A follows strategy

the possible outcomes of exchange between them can be summarised as in Table 3. The numbers in Table 3(*a*) show the increases in *A*'s satisfactions as a result of selling his apples at the exchange rate determined by any combination of his and *B*'s strategies; and similarly for *B* in Table 3(*b*).

If there is no collusion, and if neither *A* nor *B* knows which strategy his rival will choose until he is actually implementing it, then *A* will follow strategy A_3 and *B* will decide upon B_3. If *A*, for example, should begin by implementing A_1, *B* will choose B_3, for of the three strategies open to him — namely, B_1, B_2, and B_3 — it is the last which promises the greatest increase in satisfaction, given *A*'s behaviour. When *A* sees that *B* is following B_3, however, *A* will choose A_3, for given *B*'s actions that promises him (*A*) the greatest addition to his utility. The combination of strategies (A_3, B_3) will be a stable equilibrium, and it will not be altered if *A* or *B* discovers the strategy which his rival is planning to follow. If we allow the possibility of collusion, however, the outcome will not be (A_3, B_3) which promises a joint gain of only 17 utils, but rather (A_1, B_3) which promises a joint increase in utility of 19 utils, for with the latter combination of strategies each can be better-off than with the former.

The incentive to collude will only be absent if the joint gain is constant irrespective of the choice of strategies — that is, if *A* and *B* are playing a 'zero-sum' game. This would occur in our example if *A* and *B* were in all respects physiologically and psychologically identical, for their indifference maps would then be identical also. The numbers in Tables 3(*a*) and 4(*a*) are the same, but those in Table 4(*b*) have been adjusted to ensure a constant joint gain of 15 utils. In this example, the outcome will still be (A_3, B_3), and this will be reached irrespective of the initial choice of strategy by either *A* or *B*: thus, if *A* chooses A_1, *B* will follow with B_3, and observing this, *A* will respond with A_3, and given A_3 there is no strategy open to *B* which promises a greater gain than does B_3. This example yields a unique and determinate solution because it possesses a 'saddle-point'.* If we sub-

3 and *B* follows 1, then the equilibrium denoted by *La* in Diagram 120 will emerge. The numbers in Table 3 are roughly consistent with this specification of the strategies.

* The saddle-point is illustrated in Table 4(*a*) by the square (A_3, B_3). It has two properties: first, it is the highest of the row minima, and second, it is the lowest of the column maxima. For a fuller description, illustrated by tables, see Hurwicz, *loc. cit.*, p. 914.

stitute 8 for 4 in square (A_2, B_2) and 6½ for 8 in (A_3, B_2) of Table 4(*a*), the saddle-point disappears and there will be no determinate solution: now, if A follows A_1, B will choose B_3, and A will respond with A_3, B with B_2, A with A_2, B with B_3, and so on. In these circumstances, the indeterminacy can be removed by assuming that A and B each follows a 'mixed' strategy — that is, each leaves the choice of his strategy to chance, so that each is attempting to maximise the mathematical expectation of his profits rather than the profits themselves.

Thus far, we have said little about the theory of games. We have rather, described a device which conveniently illustrates and summarises the results of some of the models in Chapter 11. The tables merely summarise the range of possible outcomes. And if we possessed all the knowledge that is needed before they can be assembled, we would probably have obtained already a solution from the conventional models. These tables, however, show the first advantage of the theory of games over our previous analysis. In Chapter 11, in each model of duopoly or bilateral monopoly, we assumed that A's (or B's) behaviour was based on some expectation about B's (or A's) actions. Each such model yields one number in each of our tables: thus, if we suppose that A is the follower and B the leader, we obtain the numbers in square (A_1, B_3) in Tables 3(*a*) and 3(*b*). In a sense, therefore, our previous analysis merely helps us to delineate the range of possible outcomes, and that of the present section indicates which of these will actually emerge. When explaining the outcome in Table 3, the argument was implicitly dynamic and similar to the analysis of the simple Cournot model in Chapter 11. We could have reached the equilibrium denoted by (A_3, B_3) more quickly, however, by arguing that A will choose A_3 because it promises him higher profits that any other strategy should B follow the strategy which is least favourable to A. The assumption that each duopolist, for example, seeks the outcome that will be most favourable to him given the least favourable reaction by his rival, provides an alternative interpretation of rational behaviour. This new interpretation is probably more appropriate than the usual one of simple maximisation based on firm anticipations of rival's reactions. The second advantage of the theory of games is that it promises a more complete theory of coalitions — a theory that will include the possibility of

threats, bribes, compensations and bluff. The need for such a theory is amply demonstrated by the inadequacy of our analysis of collusive oligopoly in Chapter 11.

The theory of games has two main disadvantages in the field of value theory — one temporary and the other permanent. First, the theory has as yet only been elaborated for games with three or four players. In most economic problems, there are many players: for example, in monopoly, while there is only one seller, there are large numbers of buyers of the product and of sellers of factors, and there may be some participating spectators (the producers of the nearest substitutes). Secondly, the device of the 'mixed' strategy, which makes non-zero sum-games determinate, is unlikely to have any empirical counterpart in actual market situations. It is true that if *A* chooses his strategy at random, then *B* can never know what *A* will do until he does it. While random choice is an effective guarantee of secrecy, it is not necessarily the only one. Further, there may be occasions when an oligopolist will wish his rivals to know the content and purpose of his strategy, for if they do they will be the more likely to support and the less likely to thwart his efforts to lead the industry or group towards an equilibrium which promises the maximum joint profits.

INDEX

PRINTED IN GREAT BRITAIN BY ROBERT MACLEHOSE AND CO. LTD
THE UNIVERSITY PRESS, GLASGOW